Context and Decision:

Three approaches to representing preferences

by

Patrick Haralabos Desmond O'Callaghan

Thesis

Submitted to the University of Warwick

for the degree of Doctor of Philosophy

Doctor of Philosophy

Warwick Mathematics Institute

August 2011

Contents

Acknowledgments

First of all, I would like to thank the people who put the time into reading my work and advising me on the subject matter of this thesis. Especially, my primary supervisor Robert S. Mackay who believed in my abilities to research and provided much needed encouragement, enthusiasm, guidance and interest in my ideas and interests at the crucial early stages. He also played a major role in coordinating numerous administrative matters throughout the last three years and before that it was the clarity of his lectures that lead me to join the mathematics department. I am also indebted to my secondary supervisor Peter J. Hammond whose knowledge of the literature and innovative approach to decision theory have been respectively instrumental and inspiring. I especially thank Saul D. Jacka who stood in as my primary supervisor whilst Robert was away during my last year. It is his unwillingness to believe anything I said or wrote until it was stated clearly and correctly, and his unrivalled attention to detail in our weekly meetings that has made all the difference. I thank my PhD thesis examiners Horst Zank and Simon French for reading the original, unpolished thesis, providing me with valuable feedback and encouragement to carry on in Academia. I thank Sebastian van Strien for overseeing the thesis examination and for offering good advice at various points in time. I also thank my colleague Aron Toth who provided sound advice on many occasions and who helped me polish chapter 4 in preparation for the job market.

It has been a pleasure to be a part of the University of Warwick, first at the Economics department and then at the Warwick Mathematics Institute. It has providing a truly wonderful place to work and socialise with so many inspiring, interesting and friendly people. I acknowledge the Engineering and Physical Sciences Research council (EPSRC) for funding both my fees and maintenance for three and a half years. In so far as the practical matter of administering my study at Warwick I sincerely thank the Graduate School, in particular Rory McIntyre and Martin Mik and my Directors of Graduate Studies Mark Pollicott and Dmitriy Rumynin for understanding my situation and allowing me to extend my period of study.

Most of all, I am indebted to my family and friends. In particular to my wife Aliandra Nasif for her kindness, patience and personal support, to my son Dennis P. O'Callaghan who during this time has grown from being a toddler to being a boy and more recently to my daughter Vassiliki T. O'Callaghan who was born two weeks after the date I submitted my thesis for examination (she has been a wonderful sleeper). They have all provided me with the very best form of entertainment and motivation along the way. Finally, I am ever grateful to my father James D. O'Callaghan for his encouragement, advice and high standards of achievement, to my mother Vassiliki Tsoukala for her endless support and care and for being there whenever I needed her, to my godmother Muriel Davies who provided a perfect model of how to be a good learner as well as to my extended family in Greece, so many of whom work in education and have inspired me to do the same.

This thesis is dedicated to the memory of my beloved mother, who in more ways than I can express, made this thesis possible.

Declarations

This research is wholly and solely my responsibility and, with the exception of quoted results that are accompanied by the relevant citation, all results and errors are attributable to me.

Abstract

We present three approaches to representing families of preference relations indexed by context over a set of alternatives. Our main motivation is that existing models of preferences typically assume that there is a unique order that ranks alternatives and contexts alike, that is preferences are context-free. We argue that this is often an unnecessary and unrealistic burden on the decision-maker's capacity to decide. The first chapter is a geometric approach that improves upon the model of Gilboa and Schmeidler (2003) and case-based decision theory. The second is a general, topological approach with applications to choice under uncertainty; for context preferences it emulates the classical representation of Debreu (1954) and (1964). Finally, as a practical alternative the previous two models we extend the model of Herstein and Milnor (1953) and hence von Neumann and Morgernstern (1944) to obtain a representation for unions of mixture spaces.

Chapter 1

Introduction

1.1 Context and decision: the general case

1.1.1 The role of context

It is hard to imagine a decision problem that is completely absent of context. Context may be either implicit and in the background, or explicit and central to the problem. Examples of explicit contexts that fill the literature are decision making under risk or uncertainty, where a decision-maker has certain beliefs, conjectures or simply information about what will happen in the future, and *given this knowledge* she wishes to make a decision. Examples of background contexts include the observable socio-economic and biological characteristics of the decision-maker which are the focus of models in econometrics. Such "background variables" are often slow to change and for a given decision-maker are taken as given for a particular decision problem. Yet other examples that may rapidly change include those where the decision-maker may be completely unaware of the context in which she finds herself, such as when there are subconscious psychological states of mind or physically observable (via magnetic resonance imaging say) neuronal configurations that have a direct impact upon the decision-making at hand.

A broad definition of context would therefore include all sorts of context: conscious and subconscious, observable and otherwise, as well as both central and ancillary. Indeed at first glance, there is no reason why context might not simply refer to a complete description of the universe including its quantum states. Of course a modeler and indeed a decision-maker will tend to exclude from the decision problem matters that are either superfluous or beyond the control of the decision maker. However, as with the analysis of markets, holding the prices of all other markets fixed is only a partial analysis of equilibrium, we believe that the equivalent of a "general equilibrium" analysis of a decision problem requires identifying the space of contexts that may vary within the decision problem and impact upon the outcome.

This leads to the fundamental question of whether there are any limitations to what we may call a space of contexts for the purposes of a decision problem. For the benefit of those that are familiar with the subject matter, an affirmative answer is provided in chapter 3 in so far as we wish to consider preferences that are represented by a utility function that is continuous on the context space. These concepts will of course be defined below.

As with other fields, it serves well to first gain a good understanding of how the present model works by considering a straightforward, canonical context space. That is, consider the set of all possible probability distributions on a given set of states of the world. A particularly simple yet fruitful example of which is the following.

Example 1.1 (Based on exercise 6.B.4 of Mas-Colell, Whinston and Green (1995) [MWG]. *p208] Consider a decision maker, Val, who faces the future threat of a flood. At the time when the threat becomes imminent, two states of nature are anticipated to be of concern: "flood" and "no-flood". Suppose that at such a time, Val will know the chances of each state obtaining. The response will be one of the following two courses of action: "evacuate" or "do nothing". At present, Val seeks to define a complete, contingent plan of action. Crucially, the contingencies she considers are the set of possible probability distributions over states that she may face when the threat becomes imminent, not the states that may subsequently obtain. In particular, given there are two states, the present context space may be identified with the one dimensional simplex* $\{p \in \mathcal{R} : 0 \leqslant p \leqslant 1\}$ *where* p *is the probability of a flood.*

This canonical class of context spaces is the one that is considered in the paper "A derivation of expected utility maximization in the context of a game" by Gilboa and Schmeidler (2003) [GS03]; a paper we will frequently refer to below. Indeed, the majority of the results in this thesis are most readily understood by considering a context space that is either a one or two dimensional simplex of probability measures: respectively the case when there are either 2 or 3 states. (We formally define a simplex of probability distributions and state some useful results concerning the ambient space in which we can consider them in the appendix of this chapter.)

Another kind of context space that has been studied in the literature is a

(Cartesian) product of the non-negative integers, where a context is given the interpretation of a database or memory (the integers signify the frequency with which a case has been observed. This is the setting for the model of case-based decision-making by Gilboa and Schmeidler (1995, 2002), a model that is mathematically and methodologically close to both [GS03] and the present exposition. Indeed, the present work is intended to provide further motivation for and some substantive generalizations of the existing work of Gilboa and Schmeidler amongst others on role of context in decision theory.

Before discussing the directly relevant literature further we will introduce the essential concepts of preferences and utility in the form that is most commonly found in the literature and motivate the concept of context preferences.

1.1.2 Preferences and utility: the classical approach

It is standard in decision theory to characterize preferences with a relation $\succ$ (or $\succsim$) on some set, which is here denoted by D. The *preference relation* has the property that for all a and b in D, $a \succ b$ if and only if the decision-maker is willing to make the following statement: "a is strictly preferred to b". Mathematically $\succ$ is simply subset of $D \times D$. The alternative approach is to define a *choice function* on the set of subsets of D. The advantage of this approach is that the choice function is, in principle at least, observable from the choice behaviour of the agent. The advantage of the latter, is that it can be understood as the psychological preference of the individual and is perhaps because in general, the set D need not be actual alternatives available to the decision-maker.

For instance, a decision-maker may be willing to state “I would strictly prefer that candidate x beats candidate y in the next elections.” Such statements may be represented by a preference relation in the following way: define a to be the statement “candidate x beats candidate y in the next elections”, and b to be the statement with x and y interchanged. Of course the decision-maker can “choose” a over b, but it seems more natural to say that a is strictly preferred to b. We adopt the view that the model should allow for this possibility whilst simultaneously striving not to extend the domain of preferences beyond what is absolutely necessary, for doing so may fundamentally alter the nature of the decision problem. We return to this point below.

In our common use of language, it is understood that by making the statement that a is strictly preferred to b, the decision-maker also means b is not strictly preferred to a. A condition on the preference relation $\succ$ called asymmetry that captures this idea is found either implicitly or explicitly–see for instance Fishburn (1979) and Kreps (1988))–in almost all models of preferences. When we impose such conditions (more often referred to as axioms) on preferences, we inevitably exclude entire classes of preferences from the subsequent analysis in the hope that we may say more about the smaller class that remains.

Whilst preference relations and the conditions we impose upon them are useful in their own right as we can evaluate and test them directly, they also provide a stepping stone to a myriad of mathematical tool-kits. In particular, it is common to characterize preferences by a real-valued function, f mapping D into $\mathcal{R}$ such that $a \succ b$ if and only if $f(a) > f(b)$, and call this a utility function or utility representation. This characterization then allows us to model and analyze our decision-maker using tools from calculus and probability theory, to derive other functions (eg. demand functions), to aggregate and look at the economic behaviour of markets and economies in ways that are otherwise impossible or the very least less tractable. Importantly, it is the conditions we impose on preferences that determine the kind of function we can use to represent an agent.

It is worthwhile bearing in mind that long before preferences explicitly modeled in decision theory, utility and the tools of calculus and probability were used to study one of the oldest challenges in decision theory: a resolution to the famous St. Petersburg paradox. Indeed, Daniel Bernoulli's solution using an expected utility function (with diminishing utility of wealth) preceded the axiomatic foundations obtained by von Neumann and Morgernstern (1944) [vNM] by over two hundred years. Thus it is often the case that the conditions we impose on preferences are driven by the desire to give an already successful utility model an appropriate decision theoretic foundation.

This especially popular model of expected utility identifies the utility function up to a positive affine transformation (also known as a "cardinal scale"). That is two utility functions u, v characterize the same preferences provided $v = \kappa + \lambda u$ for $\kappa \in \mathcal{R}$ and $\lambda > 0$. This scale is the same as that with which temperature was measured before knowledge of the existence of an absolute zero . Indeed [vNM] provides a detailed and convincing discussion of the historical and physical parallels between attempts to accurately measure temperature and attempts to measure utility.

They propose that if an individual strictly prefers “a glass of tea” to “no drink” and “a cup of coffee” to “a glass of tea”,

$$u(\text{“no drink”}) < u(\text{“glass of tea”}) < u(\text{“cup of coffee”}),$$

then, letting $u(\text{“no drink”}) = \kappa$, there should exist a number $0 < p < 1$, such that

$$p\, u(\text{“cup of coffee”}) + (1-p)\kappa = u(\text{“glass of tea”})$$

or equivalently $p\,(u(\text{“cup of coffee”}) - \kappa) = u(\text{“glass of tea”}) - \kappa$. From this they conclude that the ratio of these two differences is equal to p and utility may thus be measured provided we define preferences on lotteries rather than the underlying objects and use the convexity properties of the space of lotteries as a yard stick.[1]

Whilst it is easy to see how this approach might fail (for instance there will presumably be a range of p where the individual has no strict preference in either direction between the lottery and the sure thing), there is a basic message that was novel in the 1940s: it is possible to make more precise measurements of preference than is the case if preferences only satisfy the conditions for an ordinal utility representation. For such preferences, v and u are equivalent representations whenever $v = f \circ u$ for some $f : \mathcal{R} \to \mathcal{R}$ strictly increasing ($s < t$ implies $f(s) < f(t)$), and in this case the sum that forms the above expected utility is completely arbitrary and meaningless. Indeed, even strength

[1]In this sense the above tea and coffee example is a slight abuse of notation: $u(x)$ should really be $u(\delta_x)$, where δ_x is the probability distribution that assigns mass one to outcome x.

of preference measurements are implied by the [vNM] model–for if $p > \frac{1}{2}$ in the tea-coffee example, this indicates a stronger preference for tea over nothing than coffee over tea, whilst the converse is true if $p < \frac{1}{2}$.

Herstein and Milnor (1953) [HM] observed that to measure utility up to a cardinal scale, it is not necessary to define preferences on a simplex of probability distributions. Rather, one needs preferences to be defined on a "mixture space". (We define this space and work with it in chapter 4, whilst a more detailed discussion of expected utility is provided in section (1.3.2).) In seeking a foundation for measurement theory Krantz, Luce, Suppes and Tversky (1971) [KLST] obtained alternative foundations for expected utility amongst a variety of other models. This approach places specific conditions on the set on which preferences are defined that generalize the convexity property that sets such as a simplex of probability measures satisfy. Instead, they require richness (also known as solvability) conditions. The weakest of these implies amongst other things that for a finite set the elements must be equally spaced. A set that satisfies these conditions is for instance the numbers $1, 2, \ldots, 10$.

The plethora of criticisms of expected utility theory that flourished in the literature on decision theory and eventually beyond started with the paradox of Allais (1953). (We discuss this paradox in more detail in chapter 4 where a possible resolution is offered.) Inspired by the apparent shortcomings of expected utility, mathematical psychologists such as Kahneman and Tversky (1979) [KT] introduced concepts from empirical literature on the psychology of behaviour to decision theory, thereby unifying hitherto largely separate and distinct literatures.

Prospect theory [KT] and its later version, cumulative prospect theory, by the same authors (1992) is perhaps the first model of decision-making to give a central role to context. They seek to explain, for instance, how an investor who has made losses and who has yet to come to terms with this fact will behave in a different way to one who is investing for the first time. The concept that plays the role of context in their model is the *reference point* which is usually taken to be some value of wealth. The idea is that preferences over changes in wealth give rise to a representation that amongst other things behaves differently either side of the reference point. Thus, relative to the new investor, the investor who has incurred a loss will behave in a risk seeking manner in order to recoup the losses simply because she is far below her reference point.

However, the preference model that underpins the function [KT] proposed to represent the agent entails minor changes to a result from [KLST] and must therefore hold for a fixed reference point. This is also pointed out by Schmidt (2003) and Bleichrodt (2007) en route to addressing this particular drawback. That is to say their model is silent on how the preference relation, and hence utility, varies with the reference point. Note that the authors are clearly aware of the importance of this question; as is apparent in the following quote from [KT] p.277, which also appears in Schmidt (2003):

> The emphasis on changes as the carriers of value should not be taken to imply that the value of a particular change is independent of initial position. Strictly speaking, value should be treated as a

> function in two arguments: the asset position that serves as reference point, and the magnitude of the change (positive or negative) from that reference point. An individuals attitude towards money, say, could be described by a book, where each page presents the value function for changes at a particular asset position. Clearly, the value functions described on different pages are not identical: they are likely to become more linear with increases in assets.

In the above models, and the vast majority of others, a single preference relation is defined over alternatives and treated as the primitive of the model. The situation is less clearcut for dynamic choice where preferences are often modeled as evolving with time, but nonetheless, the primacy of the initial (single) preference relation is still standard (see Hammond (1976) for a just such a model of changing tastes).

We will refer to such models as *context-free* even though it is clear that, for any given model, this property must attributed to one of the following two underlying assumptions. The first is that preferences are defined on a much larger set of statements than the set of alternatives (or mixtures thereof). That is, preferences consist of all possible statements concerning context, context-alternative pairs, mixtures thereof etc, etc. The second is that the context is simply suppressed, and we then understand the preference relation in terms of the decision-maker defining preferences over a subset of the objects of the universe whilst holding the remainder fixed on the grounds that they will not change during the course of the decision at hand. It seems natural to call these two classes of context-free preferences *global* preferences and *local* preferences respectively.

Depending on the application, models like [vNM] may fall into either of the two categories. If, as in example (1.1), the agent is making a plan for a future set of contingencies then there are clearly a whole host of possible present contexts that, depending on which one holds at the time the plan is made, may give rise to a whole host of different plans. In this case [vNM] relies on the second assumption. On the other hand, if we assume the agent's present context is identified with an element of the set on which preferences are defined, then provided the domain of preferences is suitably large, preferences are globally defined. The agent's actual decision will still depend on the present context, but preferences do not. Indeed, a special case of this situation is typical in game theory. (The discussion of the [vNM] model section (1.3.2) in section below is related to this point.)

This plurality of possibilities is not possible for instance with the Savage (1972; original edition 1954) model of subjective expected utility.[2] This is because the present context is identified (implicitly) to be the subjective knowledge that gives rise, through the structure of the domain of preferences and the conditions thereon, to the probability distribution on the set of future states that appears in the utility representation. For a different context, a different probability distribution would appear in the representation, for in this model utility itself is dependent only on the outcomes. To us it seems that this is the reason Savage (and for different reasons Binmore (2009)) suggests his model

[2] We do not provide an exposition of this model here. The reader is referred to either Fishburn (1979) or Kreps (1988).

only applies to "small worlds". Since there is only a single preference relation to which the conditions apply, the model is silent on how preferences might vary with subjective knowledge, even though the representation suggests that the subjective probability distribution might fully capture this variation in context. Indeed, even if the current subjective knowledge was somehow included in the domain of preferences, the latter would still be defined from that vantage point. Thus the Savage model of expected utility employs local preferences.

The question of why Savage chose to suppress context is important for the present discussion. Indeed, in the model of Savage, there is a another form of context that is modeled, but only indirectly. Being a model of choice under uncertainty about future states, with a single preference relation defined on the set of "acts" (the set of *all* functions from states to consequences), Savage isolates functions that differ only at a given state or event and is hence able to impose conditions on state-preferences. The rationale for doing so is to relate current preferences to what preferences would be given the knowledge that a future state or event will obtain. By taking this approach, Savage chose to substantially extend the domain of preferences rather than work with context preferences. As we will now argue, this choice is typical in the classical decision theory literature.

Although the following passage from Mertens (2003; first version 1987) does not address the literature on decision theory per se, it provides one of the most candid articulations of a key motivation of classical decision theory. In this paper, Mertens, driven by the proliferation of new equilibrium solutions in game theory that had come about through “pressure from applications”, has resolved to put game theory back on normatively solid ground.

> It is a challenging task to put those new developments on as solid a foundation as classical theory, and to show that all those intuitive and context-dependent arguments can be rationalised in a purely decision theoretic, context independent way...

Then, in referring to a proposal to address this challenge, Mertens continues:

> This concept has a number of satisfactory properties, and yields typically in the above mentioned applications the same equilibria we would have selected by the intuitive, context depending arguments.... It follows that all those new developments in economic theory are at least mutually consistent, and based on essential decision theoretic aspects i.e., not only is it true that we did not select different solutions in different contexts for games that were ordinally identical, but there is a single, fully ordinal theory, that has reasonable properties over the space of all games, and that underpins all those developments.

This quote epitomizes the normative motivation for the classical approach to modeling decision-makers and the apparent aversion to “non-ordinal” foundations and context-dependence. This aversion is perhaps due to Samuelson (1938) and the revealed preference approach to decision theory he helped found. The idea that preferences may be revealed through actual choices is extremely appealing, and the difficulty with what Mertens might call non-ordinal features of a decision model is that they can surely never be revealed or verified by observing the choices of an individual. By contrast, the context rich empirical data and tools available today are very different to those available at the time when the revealed preference approach was initiated, and they ensure that observations on choice are only one part of the set of observations.

For instance, today econometricians have an abundance of data on the individual characteristics of decision-makers, and indeed do try to explain choices conditional upon observable background data. Indeed in models of discrete choice, the objective is to try and obtain estimates of the parameters that determine how preferences change across a population that is sampled from a range of contexts. One interpretation of this methodology is that, up to the unobservable context (which is modeled as noise) it is the observable context that is determining decisions. Indeed the approach suggests that, noise aside, this is what we would all do if we were in the same (observable) context.[3]

Today it is common for empirical psychologists and experimental economists to aggregate data across individuals and seek to measure the typical response to a given stimulus. Often, the agent will be unaware of the stimulus, and in such circumstances it is difficult to imagine that the decision-maker is ranking not only the alternatives before her, but also the stimuli. Thus the fact that there is a behavioural response to stimulus will, in general, tell us nothing about whether a given agent has preferences for one kind of stimulus or another. Similarly, neuroscientists have the possibility of scanning neural ac-

[3]This interpretation adheres to the view attributed to Hume (1739). In his essay on the philosophical question of liberty versus necessity (also known as determinism versus free-will) Hume convincingly argues that free will is the illusion the decision-maker experiences by living the decision, that in practice, any other decision-maker in precisely the same situation or context (molecules, atoms and all) would have reached the same decision, by the same means. He concludes from this that determinism and free-will may co-exist, with free-will being a purely subjective notion and the determinism a purely objective notion.

tivity, thereby gathering information that may be non-ordinal (and thus contextual) by its very nature.

Similar to the way expected utility was used long before [vNM], it appears as if the preference foundations are lagging behind the applications. We believe that there is therefore much room for modern decision theory to cautiously embrace context-dependence, and moreover to do so on a firm theoretical basis. One of the objectives of the present thesis is to help lay the foundations for such a transition.

1.1.3 Context preferences

With the arguable exception of the literature on dynamic choice, models in decision theory that take as primitive a family of preference relations on some set and indexed by another rare. Such models are more commonly associated with the literature on social choice, where a collection of agents are modeled as a collection of preference relations. Indeed, the decision theory model of [GS03] which is closest to the present work and in particular to the model of chapter 2 is mathematically similar to the work of Young (1975), Myerson (1995), Ashkenazy and Lehrer (2001) and Azrieli (2011) all of which form part of the literature on social choice.

In general context preferences are defined in the following way.[4] For a given set of contexts X, upon which preferences are *not* defined and another set D,

[4]Our use of the term "context preferences" as opposed to "context-dependent preferences" stems from the more common use in language of similar abbreviations such as "university students" for the more accurate but tedious "university-attending students". Indeed at times, following Schmidt, we will abbreviate further still to just "preferences".

we define context preferences to be

$$\{\succ_x \subset D \times D : x \in X\}. \tag{1.1}$$

For a particular context x we will refer to $\succ_x$ as the preference relation at x. Note that at this stage, prior to imposing any conditions on preferences, this notation is general enough to include the case where for some x and y in X, the relations $\succ_x$ and $\succ_y$ are defined on D_x and D_y respectively where $D_x \neq D_y$. This is because we may define $D := \bigcup_{z \in X} D_z$, and both $\succ_x$ and $\succ_y$ are subsets of $D \times D$.

In the above definition, the set D is not identified with the set of alternatives A, for in general, a context preference relation $\succ_\cdot$ (this notation is intended to denote a general $\succ_x$ for some $x \in X$) may contain not only certain pairs of alternatives but also pairs of contexts, pairs of alternative-context pairs, and even higher order concepts. Reconsider the flood example, where the context space is the set possible chances of a flood. We would surely agree that the statement "I strictly prefer the situation where there the chances of a flood are roughly $\frac{1}{4}$ to that where it is roughly $\frac{3}{4}$" will be a part of the typical agent's preferences.

Should statements of the form "I would prefer *to evacuate when the chances of a flood are* $\frac{1}{50}$ over *staying at home when the chances of a flood are* $\frac{1}{10}$" be assumed to be a part of the agent's preferences? If the agent is willing to make such statements, then the domain of preferences, D, should perhaps be extended from the set A of alternatives to the set $A \times [0, 1]$ of alternative context pairs. What the set X will then be is dependent upon the problem at hand. We consider this kind of extension in chapter 4.

The case of prospect theory, where a level of wealth is often taken to be the context, the situation is similar to the one just described: it seems somewhat unnatural to altogether exclude statements such as "I strictly prefer more wealth to less" from the set of preference statements of an agent. This seems a fair criticism of Schmidt (2003). He defines context preferences as follows: the set D is taken to be the set of probability distributions on $\mathcal{R}$ that have finite support (assign positive measure to a finite number of points); the elements in the support of the probability distributions are understood to be changes in wealth; finally, the set of contexts is composed of the possible levels of wealth the agent might treat as her reference point and this too is equated with set of real numbers. The context preferences are of the same form as (1.1).

Indeed the same criticism may be leveled at [GS03], where the context space is probability distributions as well as the theory of case-based decisions where it is memories, for surely there are memories that the agent likes more than others. On the other hand, it is also likely that an agent's preferences will be incomplete over the set of memories. (Incomplete context-free preferences arise when the weak preference relation $\gtrsim$ is not complete. That is, there exists a and b in D such that neither $a \gtrsim b$ nor $b \gtrsim a$.) Indeed, the context preferences approach can be seen as a constructive and structured approach ("bottom up") to dealing with incomplete context-free preferences.

Although we will not pursue the next issue any further in this thesis, it is an obvious direction for future research on context preferences. Suppose a modeler is seeking a comprehensive representation of the agent's preferences (for the purposes of a general equilibrium type analysis say). One approach is to identify a set X_c of statements about the world for which, the agent is able to say that "conditional upon the truth of [any such statement], my preferences are …". Call this the set of contexts about which the agent is conscious.[5] The remaining contexts, of which the agent is either unaware or unable to distinguish as a context in its own right, but which the modeler deems to be relevant to the decision, might be denoted as X_u.[6] The Cartesian product of these two sets would then define the context space X.

Due to time constraints and in order to facilitate easy comparison with the existing literature, the present thesis will work in the same setting as, Schmidt (2003), Bleichrodt (2006), and in particular [GS03]. Indeed as they do we will identify the set D with the set of alternatives, and with the exception of the first part of chapter 3 and chapter 4, we will let X be the set of probability distributions over a given set of states. We discuss the notation in the next section.

Utility representation of context preferences

A utility representation of context preferences is a family of classical utility functions, one for each context. Thus, a utility representation or characterization of context preferences is a function $f : X \times D \to \mathcal{R}$ such that for any a, b

[5] This approach appears to fit well with the literature on *epistemic logic*–see Kaneko (2002) for an introduction to this literature and its use in game theory.

[6] As we have discussed above, examples of this consist of psychological states or neuronal configurations about which the agent is completely unaware.

in D and x in X we have

$$a \succ_x b \quad \text{iff} \quad f(a,x) > f(b,x) \tag{1.2}$$

If for a given representation f, $f(\cdot,x) > f(\cdot,y)$ for some x and y in X, this ought to be incidental and particular to f. That is, we should be able to find another function g satisfying (1.2) such that $g(\cdot,y) > g(\cdot,x)$. So whilst the main preoccupation will concern the existence of a utility representation, the extent to which the representation is unique is still central to the discussion, just as it is in the setting of [vNM] and [HM].

Without extending preferences to include contexts, utility representations of context preferences contain less information and are therefore unique up to a larger class of transformations than the single positive affine one that distinguishes the [vNM] and [HM] models. In the general model we present in chapter 3, far less may be said. But if as in chapter 2, heuristically speaking, the following two properties hold (i) the information that context preferences provide is rich enough, and (ii) the representation preserves the natural mathematical operation for the given space of contexts (eg. mixtures on convex set; positive combinations on the positive integers), then utility will be *cardinally measurable and unit comparable* (CUC)–see the appendix of the present chapter for a detailed discussion of measurability and comparability.

Conditions across contexts

Once the context space and domain of preferences has been identified, the next task is to identify suitable conditions that may be imposed on the resulting family of preference relations. Conditions such as asymmetry (discussed above for the case of context-free preferences), may of course be imposed upon any given preference relation $\succ_x$, for any context x. In contrast to the context-free case, conditions may also be imposed that restrict how preferences vary across contexts. We may refer to this as an *inter-context* condition.

An important example of such a condition is the "combination" condition of [GS03]. The more natural part of this condition states that for any a, b in D, if $a \succsim_x b$ and $a \succsim_y b$, then for every context z that is a strict convex combination of x and y, $a \succsim_z b$. This condition is clearly enough to ensure that the set of contexts x for which $a \succsim_x b$ is convex. In turn, this is necessary for the existence of a utility function that is mixture preserving across the context space that [GS03] work with.[7]

It is worth mentioning that the *sure-thing principle* of Savage (1954) is, in essence, an inter-context condition. As we have discussed above, Savage, using context-free preferences defined on the set of all functions from states to outcomes, is able to indirectly impose consistency conditions that relate preferences given current knowledge to preferences conditional upon events obtaining. He introduces the condition as follows:

> If the person would not prefer f to g, either knowing that the event B obtained, or knowing that $\sim B$ *[interpreted as "not B"]* obtained, then he does not prefer f to g. Moreover, (provided he does not regard B as virtually impossible) if he would definitely prefer g to

[7] That space is the convex set of (finitely additive) probability measures on a measurable state-space.

> f, knowing that B obtained, and, if he would not prefer f to g, knowing B did not obtain, then he definitely prefers g to f.

In other words, this condition implies firstly that if, given any event has obtained, g is weakly preferred to f, then it is also weakly preferred given the current information. Secondly, if this is true for all events and for some event preference is strict, then the first act is strictly preferred given the current information. Skiadas (1997) adopting the context preferences approach, holding the current information fixed throughout, formulates the condition in this way. Note that the first part of Savage's sure-thing principle is similar to what we called the more natural part of the [GS03] convexity condition above. What is not often appreciated is the strength of the second part. A similarly strong form of "breaking the tie" constitutes the second part of [GS03]. We will look at the implications of relaxing this condition in the second part of chapter 3.

1.2 Context and decision under uncertainty

1.2.1 Notation for the canonical setting

As we have already mentioned, in what follows, with the exception of the first part of chapter 3 and chapter 4, we will work with a canonical context space that is the set of probability distributions over a set of states and for each context, define a preference relation over the set of alternatives. Thus the canonical model applies to decisions under risk or uncertainty.

Let $S := \{s, t, u, \ldots\}$, $1 < |S| = n$, denote an abstract set of future states of nature, and $\Delta \equiv \Delta(S)$ the set of probability distributions over S with elements denoted by p, q, and r. For all $p \in \Delta$, let $A := \{a, b, c, \dots\}$, $1 < |A| = m$, be the set of alternatives on which the agent's strict preference relation, $\succ_p$, is defined.[8] For a given p, $\succ_p$ is a binary relation over A. The following notation for a family of preference relations

$$\{\succ_p \subset A \times A : p \in \Delta\}, \quad \{(A, \succ_p) : p \in \Delta\}, \quad \text{and} \quad (A, \succ_p)_{p \in \Delta},$$

will refer to the same object: the context preferences or, simply, preferences of the individual agent.

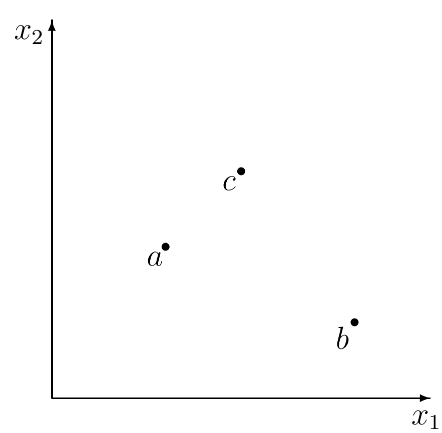
Each strict preference relation is therefore a subset of the collection ordered pairs of elements of A satisfying: *at context p, for any alternatives a and b in A, a is strictly preferred to b if and only if $a \succ_p b$.* When no strict preference holds in either direction for some pair (a, b) (that is neither $a \succ_p b$ nor $b \succ_p a$), we write $a \sim_p b$. Note that by definition, $\sim_p$ is reflexive and symmetric, but not necessarily an equivalence relation as transitivity ($a \sim_p b$ and $b \sim_p c$ together imply $a \sim_p c$)may fail to hold.

This notation (for context-free preferences) is used by both Fishburn (1970) and Kreps (1988) in their classic textbooks on decision theory. Our choice of strict as opposed to weak preference as primitive is partly due the appeal of the asymmetry condition discussed in section (1.1.2) and partly due to the fact that a statement of strict preference by an agent is more concrete and readily observable than its weak counterpart.

[8] Unless otherwise stated, the sets A and S will be understood to be countable. Also, note that the set of alternatives are not state-dependent because they are available to the decision-maker prior to any particular state obtaining.

This approach also allows us to distinguish transitivity of $>_.$ from that of $\sim_.$. The latter is widely recognized as failing to hold in simple examples: consider an agent who strictly prefers no sugar at all to a single sugar cube in her glass of tea, but who is unable to distinguish between two glasses of tea, one of which contains a single granule more than the other. Transitivity of strict preference appears much more natural and mathematically useful than than that of $\sim_.$ which may also capture more general kinds of incomparability as the following example shows.

Example 1.2. *Let A be a collection of points in $\mathcal{R}^2$ and let $>_.$ denote a preference relation over A corresponding to some context for which we suppress notation. If $>_.$ coincides with the strict order on $\mathcal{R}^2$, then $x >_. y$ iff both $x_1 > y_1$ and $x_2 > y_2$ are true (where $>$ is the standard order on $\mathcal{R}$ and the subscripts on x and y denote the first and second axes).That is to say, c is strictly preferred to a if and only if it lies strictly to the north-east of a. Clearly, by transitivity of $>$ on $\mathcal{R}$, $>_.$ is transitive.*



Any pair of points that lie to the south-east or north-west of one another are incomparable under this order. Thus in the above diagram, a and b are incom-

parable under $\succ_.$, as are b and c. In such cases, as no strict preference holds in either direction, and we may write $a \sim_. b \sim_. c$. However since $c \succ_. a$, we see that $\sim_.$ fails to be transitive.

Although we do not relax the condition that $\sim_.$ be transitive in the present thesis, our approach is the natural starting point from which to do so.

Set notation for context space

For any subset E of A, the set of p in Δ satisfying $a \succ_p b$ for every $b \in E$ is denoted by $\mathcal{B}_a(E)$, thus a is (strictly) better than every element of E for all $p \in \mathcal{B}_a(E)$. Similarly, for any subset E of A, the set of p in Δ for which b is strictly preferred to a for all $b \in E$ is denoted $\mathcal{W}_a(E)$, so that a is (strictly) worse than every element of E for each $p \in \mathcal{W}(E)$. We will make extensive use of the shorthand $\mathcal{B}_{ab}$ for $B_a(\{b\})$ and similarly $\mathcal{W}_{ab} := \mathcal{W}_a(\{b\}) \equiv \mathcal{B}_b(\{a\}) =: \mathcal{B}_{ba}$. Finally, the set of $p \in \Delta$ for which $a \sim_p b$ will be called $\mathcal{N}_{ab}$, and the set $\mathcal{N}_{abc}$ is defined as $\{p \in \Delta : a \sim_p b \sim_p c \sim_p a\}$.

We make use of the fact that S has n elements to identify Δ with the $n-1$ dimensional simplex in $\mathcal{R}^n$, $\{p \in \mathcal{R}^n_+ : p_1 + \cdots + p_n = 1\}$, where $\mathcal{R}^n_+$ is the set of vectors in $\mathcal{R}^n$ for which every element of the vector is strictly positive, and $\bar{\mathcal{R}}^n_+$ will denote the closure of this set in $\mathcal{R}^n$, that is

$$\bar{\mathcal{R}}^n_+ := \{p \in \mathcal{R}^n : p_i \geqslant 0 \text{ for all } i = 1, \ldots, n\}.^9$$

[9] We prefer this notation on the grounds that it is more in line with modern notation used in the mathematical community outside economics. The standard notation in economics being $\mathcal{R}^n_{++}$ for the strictly positive vectors and $\mathcal{R}^n_+$ for the non-negative vectors.

Where useful, we will consider Δ as a subset of $\bar{\mathcal{R}}_+^{n-1}$, that is we will identify each probability distribution in Δ with a point in the set $\{p \in \bar{\mathcal{R}}_+^{n-1} : p_1 + \cdots + p_{n-1} \leqslant 1\}$. In appendix (1.B) of this chapter, we provide some essential results that justify the relationships between these spaces.

1.2.2 Motivation for the canonical setting

In this thesis we will look at how and when it is possible to use a real-valued function to characterize and represent the preferences over a set of alternatives of a decision-maker facing variety of contexts in which the decision problem may arise. In the canonical decision problems upon which the following exposition predominantly focuses, one way to view the decision problem is in terms of a decision-maker with the following knowledge:

i. she knows she *will* face a decision in the presence of uncertainty about the future state of nature;

ii. she knows, that when the situation arises, she will have sufficient information to be able to identify the likelihood of any given state of nature;

iii. given any probability distribution over states, she knows whether or not she has a strict preference for one course of action, in response to the uncertainty, over another.

The flood example we have discussed above is an instance of this situation. We also introduce the next example where there are three states and three alternatives.

Example 1.3 (Getting to and from University)**.** *Val lives near her university campus and each day she decides whether to walk, catch the bus or cycle to* *university. Thus* $A := \{a, b, c\}$ *with a:="walk", b:="bus" and c:="cycle". Suppose that the following conditions hold:*

(i) bicycles are not allowed on the bus;

(ii) the bike may be stolen if it is left on campus overnight; and

(iii) the cost of return bus ticket is significantly cheaper than two singles.

These conditions contribute to Val's strict preference for traveling by the same means to and from university on any given day. This means that her decision to buy a bus ticket or ride the bicycle to university in the morning amounts to a commitment to a single means of travel for the whole day. Of course if she walks to university, she can catch the bus back, but if she knows that conditions will be such that in the afternoon she would certainly not want to be walking home, then she would always buy the ticket, say, in the morning.

As is typical in Autumn, the weather may vary substantially during the day, and Val knows that it may be raining, sunny or snowing/icey when she returns that day: so $S := \{r, s, t\}$ *with r:="raining", s:="sunny" and t:=snowing/icy. Before setting off each morning, the radio weather forecast gives her the chances,* $p = (p_s)_{s \in S}$*, of each of the states obtaining.*

Fed up with deciding each day, Val decides, once and for all, to devise a complete, contingent plan of what to do. At the time she makes her plan, the set of contingencies is the set Δ*. (Here* Δ *is isomorphic to the 2-dimensional, unit simplex in* $\mathcal{R}^3$*.)*

Suppose that if she were certain it would be sunny in the evening, Val would strictly prefer walking to cycling to going by bus, ie. $a \succ_s c \succ_s b$.[10] *When it will rain for sure, her preferences are* $c \succ_r b \succ_r a$ *and when it will snow or be icey for sure, her preferences are* $b \succ_t a \succ_t c$.

It seems plausible that Val's preferences conditional upon p, $\{(A, \succ_p) : p \in \Delta\}$, *will vary with the probability she face on any given day: the question is how. Suppose that at the probability* $p = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, *Val's strict preferences are the same as they are when state* t *occurs for sure. Is it then plausible for instance that, for all probabilities* $q(\lambda)$ *such that* $q(\lambda) = \lambda\delta_t + (1-\lambda)p$ *and* $0 < \lambda < 1$, *Val's strict preferences are unchanged? That is, if we start at a given context* $q(0) \equiv p$ *and increase the probability of state* t *in such a way that the ratio of the probabilities of state* r *and* s *remain unchanged, does strict preference remain the same? If so then we conclude that the set* $\text{conv}(\delta_t, p)$ *is a subset of*

$$\{p \in \Delta : b \succ_t a \succ_t c\} \equiv \mathcal{B}_{ba} \cap \mathcal{B}_{ac}.$$

Whilst context preferences of the canonical form presented above are well suited to modeling the above example, which might be viewed as a game against nature where nature presumably does not respond to Val's choices: they are arguably inadequate for the analysis of strategic game rock-paper-scissors, where a player is penalized for being predictable or where preferences may be defined over correlated strategies. We discuss this matter further in chapter 4.

[10]This involves a slight abuse of notation as s is used instead of the probability measure assigning probability one to state s, δ_s.

Remark 1.1. *The time-line of the example above has three obvious stages: the final stage, after both the action has been taken and the uncertainty is resolved; the penultimate stage, when the agent must act; and the first stage, when the problem is contemplated and a complete, contingent plan of which action is to be taken in the penultimate stage is laid out by the decision-maker.*

Our focus is on behaviour in the first stage. In the language of Hammond (1981) we are taking the "ex-ante" approach to choice under uncertainty / risk. In this thesis no considerations will be given to the optimal time to act. Similarly, the related question of what constitutes "sufficient information to be able to identify the likelihood of any given state of nature", as stipulated in (ii.) above, is also avoided. That is to say, we model neither the process, nor the assumptions on behaviour that are involved in the decision-maker converting information about the uncertainty into beliefs and deciding whether or not to act.

Of the three basic assumptions on the knowledge held by the decision-maker (iii.) is the strongest and most questionable; although it is not as strong as it may at first seem. This is because the decision-maker is free to have no-strict-preference for one action over another: for every possible probability distribution over states, and such preferences would not necessarily indicate indifference.

It is a central tenet of this thesis that even when precise information regarding the uncertainty is available, choosing between courses of action may still be the most difficult part of the decision-making process. The agent must still strive to understand and weigh-up the differences between the alternatives available to her. So all other things being equal, a model in which the domain of preferences is minimal and observable to the modeler ought to dominate others.

Not enlarging the domain of preferences comes at the price of requiring that the agent is able to imagine, or recall from past experience, what her preferences are in situations different from the present context. That is, we place greater demands on her ability to imagine variations in the circumstances she may face in the second stage. Whilst there is no denying that this is the main drawback of this approach, it is a weaker assumption than to assume preferences over contexts and alternative-context pairs.

Finally, defining preferences on an observable domain such as actions rather than the much more abstract concept of outcomes or consequences, that may or may not be dependent upon states or other factors, brings the model closer to those that deal with choice functions and observed behaviour. As is the case in that literature, it may well turn out that the decision-maker behaves as if she were contemplating all the outcomes and/or consequences of her actions. If so, this will be an output of the model rather than a starting point.

1.3 Expected utility for context preferences

Given that our canonical example is a model of decision-making under risk or uncertainty, it is necessary to discuss in more detail the variety of approaches that already exist. This section provides an exposition of the different models of expected utility that may be found in the literature and shows how there is really only one candidate for context preferences.

1.3.1 The function-space approach to expected utility

A initial demonstration of why the standard approaches will not work and why new inter-context conditions are necessary is provided by the following discussion. We consider a theorem for finite alternative and state spaces by Adams (1965) and Fishburn (1979). It will show that in our setting, where a complete contingent plan for each p is the objective, the theorem of Fishburn gives a family of additive representations, one for each p. However, as the condition has nothing to say about how preferences vary with p, it is not sufficient to ensure linearity in p.

Let A_s denote the set of outcomes in state s of each of actions in A. The next theorem defines preferences on $\times_{s=1}^{n} A_s$.

Example 1.4. *In example 1.3 recall that*

$$A := \{walk, bike, bus\} \equiv \{a, b, c\}, \quad \textit{and}$$

$$S := \{rain, sun, snow/ice\} \equiv \{r, s, t\}.$$

It is not only arrays of outcomes such as (a_r, a_s, a_t) *(which may be identified with the alternative "walk" lie in* $A_r \times A_s \times A_t$*) that lie in the domain of preferences; the idealistic but unfeasible array* (c_r, a_s, b_t) *is also there. As a result even in a simple problem such as this the cardinality of the space* $A_r \times A_s \times A_t$ *is 27, whereas* $A \times S = 9$*.*

Definition 1.1 (Fishburn (1979): Equivalence relation on $\times_{s=1}^{n} A_s$)**.**

$(\alpha^1, \ldots, \alpha^J)\, E_J\, (\beta^1 \ldots \beta^J)$ *if and only if* $J > 1$, $\alpha^j, \beta^j \in \times_{s=1}^n A_s$ *for* $j = 1 \ldots, J$, *and it is true for each* s *that* $\alpha_s^1, \ldots, \alpha_s^J$ *is a permutation of* $\beta_s^1, \ldots \beta_s^J$.

For a fixed, suppressed context we have the following result.

Theorem 1.1 (Scott and Suppes (1958), Fishburn (1979))**.** *The relation* $\succ_.$ *on* $\times_{s=1}^n A_s$ *satisfies the following condition: for all* J*-sequences,* $\{\alpha^j\}$ *and* $\{\beta^j\}$ *in* $\times_{s=1}^n A_s$*, and all* $J = 2, 3, \ldots$

(F) if $(\alpha^1, \ldots, \alpha^J)\, E_J\, (\beta^1, \ldots, \beta^J)$ *and* $\neg(\beta^j \succ_. \alpha^j)$ *for* $j = 1, \ldots, J-1$*, then* $\neg(\alpha^J \succ_. \beta^J)$*,*

if and only if there exist real valued functions $u_1, \ldots, u_n$ *on* $A_1, \ldots, A_n$ *respectively such that, for all* $\alpha, \beta \in \times_{s=1}^n A_s$*,*

$$(4) \quad \alpha \succ_. \beta \quad \Leftrightarrow \quad \sum_{s \in S} \big(u_s(\alpha_s) - u_s(\beta_s)\big) > 0.$$

One problem with the condition (F) in the above theorem is that it is somewhat difficult to evaluate without case by case verification. Moreover verification is by no means an easy task: in example (1.3) where there are three alternatives and three states, in order to verify the condition, we have to make $\binom{27}{2} = 351$ comparisons.

Another problem with the above representation is that it gives us no information about how preferences vary with the context that is the current belief. Presumably, for context preferences $(\succ_p, p \in \Delta(S))$, the theorem holds for any given p, and the representation is as follows:

$$(5) \quad \alpha \succ_p \beta \quad \Leftrightarrow \quad \sum_{s \in S} \big(u_s(\alpha_s, p) - u_s(\beta_s, p)\big) > 0.$$

If we seek a representation that has more structure, and thereby resembles a plan of action we need more conditions on how preferences vary with context. Chapters 2 and 3 provide two approaches to dealing with this problem.

By contrast, the classical approach to adding more structure has been to suppress reference to p and focus on state-preferences. In the Anscombe-Aumann (1963) approach, the first step is to restrict the space of outcomes to be the same in each state and define lotteries over these outcomes.[11] The second step is to impose that the ordering of the set of outcomes in each state is the same across states. This state-independence allows us to identify the actual probability measure p that represents the information context (beliefs) of the agent at decision time. Thus the Anscombe-Aumann approach says nothing about how preferences vary with the information context, and hence p. Perhaps it is best to view this model as capturing behaviour at stage two in the time-line we describe in remark (1.1) above.

The same may be said of the models of Savage (1954), Scott and Suppes (1958), and Debreu (1959). They seem to be more appropriate for addressing the problem of an agent making a decision in the heat of the moment and at a particular information set. By contrast, the present approach is closer in this sense to [vNM] where a complete, contingent plan of action is outlined in the first stage.

[11] Or at least ensure that there exist two outcomes that are common to all states (Fishburn (1979)).

1.3.2 EU for context preferences

Let us consider the form of an expected utility (EU) function *representing* the preferences of an agent over a set A *given the knowledge of a probability distribution*, p, over the state space S.

Definition 1.2 (EU function representation of context preferences).
$U : A \times \Delta$ is an EU representation of context preferences if both the following are true:

i) there exists $u : A \to \mathcal{R}^n$ such for all p in Δ and a in A,

$$U(a,p) = \mathbf{E}_p u(a) := \int_S u(a,s)\mathrm{d}p(s) \quad \big(or \quad \sum_{s \in S} p_s u_s(a)\big);$$

ii) U is said to represent context preferences, $\{(A, \succ_p) : p \in \Delta\}$, if, for every $p \in \Delta$ and $a, b \in A$ we have

$$a \succ_p b \qquad \Leftrightarrow \qquad U(a,p) > U(a,p).$$

From the above definition, it is clear that the vector-valued function $u : A \to \mathcal{R}^n$ characterizes the representation. Indeed, the set $\mathcal{N}_{ab}$ is contained in the hyperplane perpendicular to this vector. As we describe in appendix (1.A), the elements of this vector-valued function are called state utility functions. The rationale is simply that if $p_s = 0$ for all $s \neq t$ then the decision maker would face state t with certainty, so that the expected utility $\mathbf{E}_p u = u_t$ which represents state preferences $\succ_t$. The essence of this EU representation is that preferences at p are represented by the weighted average of the state-utilities, where for each state s, the weight is defined to be p_s, the probability of state

s.

1.3.3 Incomparability across states

Suppose context preferences give rise to a representation that is cardinally measurable but non-comparable across states (CNC across states). In this case, as the following discussion shows, it is not clear whether probability plays any role at all. First we define the *support* and *null-sets* of a probability measure.

Definition 1.3 (Support of p)**.** *For countable S, the support of a probability measure p, denoted* $\text{supp}(p)$*, is the set of elements of S for which p assigns positive measure. Thus,* $\text{supp}(p) := \{s \in S : p_s > 0\}$.

Definition 1.4 (Null-sets of p)**.** *The null sets of a probability measure p, are these for which p assigns zero probability. That is, if $T \subset S$ and $p(T) = 0$, then T is a null set.*

Let preferences be represented by U and fix $p \in \Delta$. Let T be set of states such that $p_t = 0$. As preferences are CNC, any other function $V : A \times \Delta \to \mathcal{R}$ that is also and EU representation satisfies the property that there exists $v : A \to \mathcal{R}$ such that for each $s \in S$, $v_s = \lambda_s u_s + \kappa_s$ for some $\lambda_s > 0$ and $\kappa_s \in \mathcal{R}$. For all $s \in T$ let $q_s = 0$. Then for all $s \in S \backslash T$, let $0 < q_s < 1$ and let $\lambda_s := q_s/p_s$, then we have

$$\begin{aligned}
\sum_s p_s u_s(a) > \sum_s p_s u_s(b) &\Leftrightarrow \sum_s p_s \, \lambda_s\big(v_s(a) - v_s(b)\big) > 0 \\
&\Leftrightarrow \sum_s q_s\big(v_s(a) - v_s(b)\big) > 0
\end{aligned}$$

so that if the q_s sum to 1, the latter is an equivalent representation indexed by the probability q. In particular this means that for all a and b in A, the relative interior, $\mathrm{ri}\,\Delta$, of Δ is a subset of one (and only one) of the sets $\mathcal{B}_{ab}$, $\mathcal{N}_{ab}$ and $\mathcal{W}_{ab}$. This rather extreme property of preferences fact has lead many authors in the literature to suppress the probability altogether (see Fishburn (1970) Ch.4 and 5, Debreu (1959)), so that the representation takes the form $\sum_s w_s$, with $w_s := p_s u_s$ for each s. Indeed, it seems reasonable to say that $(w_s, s \in S)$ characterizes the representation of preferences $\succ_p$. This kind of representation, contrary to what one might think, is in fact overly restrictive in that it rules out whole classes of preferences without any clear reason.

1.3.4 vNM representation of context preferences

CNC preferences are in contrast to the class of state independent models of [vNM] and Anscombe and Aumann (1963). Now suppose the context preferences give rise to a [vNM] type representation on consequences $Y = A \times S$. The resulting utility function is characterized by a single utility function $u : Y \to \mathcal{R}$ and for every π and ρ in Δ_Y we have

$$\pi \succ \rho \qquad \Leftrightarrow \qquad \sum_{x \in Y} \pi_x u(x) > \sum_{x \in Y} \rho_x u(x). \tag{1.3}$$

Recall that if v also satisfies the above equivalence then we should have $v = k + lu$ for some $k \in \mathcal{R}$ and $l > 0$.That is, preferences are cardinally measurable and fully comparable (CFC–again see the appendix to this chapter for more on these concepts). As k and l do not vary with s, for given preferences, the vNM axioms define a much smaller invariance class of functions that represent preferences than is the case for CNC preferences.

In seeking to retrieve "response" or context preferences from a vNM representation,[12] we restrict attention to elements π of Δ_Y such that $\pi = p \times \mu$ for some $p \in \Delta_S$ and $\mu \in \Delta_A$ that is elements of $\Delta_S \times \Delta_A$ (where Δ_A is the set of probability distributions on A). Preferences over Δ_Y restricted to $p \times \mu$ and $p \times \nu$ in $\Delta_S \times \Delta_A$, lead to the representation characterized by $u : Y \to \mathcal{R}$ in the following way. For each b and c in A:

$$\begin{aligned} \delta_b \succ_p \delta_c \qquad &\Leftrightarrow \qquad \sum_{(a,s)\in Y} p_s u(a,s)\big(\delta_b - \delta_c\big) > 0 \\ &\Leftrightarrow \qquad \sum_{s\in S} p_s \big(u(b,s) - u(c,s)\big) > 0. \end{aligned}$$

In contrast with CNC preferences, there are now far fewer restrictions on the sets $\mathcal{B}_{ab}$, $\mathcal{N}_{ab}$ and $\mathcal{W}_{ab}$. Of course, each of these sets must be convex, and if either of $\mathcal{B}_{ab}$ and $\mathcal{W}_{ab}$ is nonempty, then $\mathcal{N}_{ab}$ is thin (empty interior relative to Δ).

Consider the special case where, for all a and b, either a dominates b for all $s \in S$ or the reverse holds for all $s \in S$, then for every a and b, convexity implies that Δ is equal to either $\mathcal{B}_{ab}$ or $\mathcal{W}_{ab}$. In this case, Δ lies in a "chamber" defined as the interior of an intersection of the $\binom{n}{2}$ open half-spaces defined by the hyperplanes $\{\mathcal{I}_{ab} : a, b \in A\}$.

This in turn implies that we may rotate, by a suitably small amount, the

[12] See Morris and Ui (2004) for the use of the terminology "best response" and "better response" equivalence classes of games. They go in the opposite direction, seeking to identify vNM preferences that "rationalize" the best and better response sets of players in a given game.

vectors $u(a) - u(b)$, keeping Δ in the corresponding chamber, to obtain a new function $v : A \times S \to \mathcal{R}$ that characterizes some expected utility representation V and such that $v \neq l\,u + k$ for any $l > 0$ and $k \in \mathcal{R}$.

Thus a vNM representation, to borrow from the language used in the social choice literature[13], *may* imply a good deal more information regarding the decision maker's preferences than can be observed from context preferences alone.

Even in the case where for every a, b in A, the sets $\mathcal{N}_{ab}$ of full dimension and the hyperplanes $\mathcal{I}_{ab}$ are fully identified, the appropriate degree of comparability for an EU representation of context preferences will turn out to be "unit comparability" and the resulting representation is referred to as CUC. It is characterized by an EU function with the property that the state utilities u_s in the vector-valued function $u : A \to \mathcal{R}^n$, are unique up to a common scale l, but independent origin k. That is, if v corresponds to another representation of the same preferences, then $v_s = l\,u_s + k_s$. We derive such a representation in chapter 2.

More generally, there appear to be no conditions one can meaningfully impose on preferences that ensure the existence of an expected utility representation of context preferences. This means that by insisting upon such a representation we are excluding perfectly reasonable preferences from the analysis without any real justification. This provides motivation for the third and fourth chapters.

[13] See d'Aspremont and Gevers (2001) or Blackorby, Donaldson and Weymark (1984) for a comprehensive survey and introduction respectively.

1.4 Synopsis

Chapter 2 presents a model that is closely related to Gilboa and Schmeidler (2003) (GS03). In this paper, the authors provide a re-interpretation of their model of *case-based decision* theory and work in the same context space as we do in this thesis. That is, where states play the role of cases and probabilities or beliefs play the role of context. We improve upon their model in two respects.

Firstly, we weaken the diversity condition in an important way. Secondly, we provide an alternative to the "combination" condition of GS03 that gives rise the linear form of the representation. The resulting conditions are weaker and we argue more intuitive.In addition, this chapter also sheds light on the role of *expected preference* representations (similar to those of Vind (1991)) for context preferences. Expected preference representations are inseparable across pairs of alternatives, but still linear in probabilities. They can be used to model intransitive preferences just as well as transitive preferences. As such it provides way of axiomatizing of the Loomes and Sugden (1982) model of intransitive regret. This is in the presence of much simpler axioms than Fishburn's skew-symmetric bilinear model.

Chapter 3 proposes a model that gives rise to a utility representation that is simply order-preserving over alternatives and continuous over contexts. We identify the most general type of context space for which we can hope to guarantee a continuous representation. The mathematics of this approach is borrowed from topology. Then we apply this model to setting of uncertainty.

In due course we intent to extend this model to provide a link between decision theory and the discrete choice models of econometrics on the one hand and neuroscience on the other.

In the fourth and final chapter we suppose that the decision-maker is willing to extend the domain of preferences to include the space of contexts. That is, the decision-maker is willing to make statements of the form: "I prefer to be in context p and choose alternative b than to be in context q and choose alternative c". Despite this extension, the domain of preferences is still much smaller than the [vNM] model of expected utility requires.

As a result, rather than a single mixture space–as defined in Herstein and Milnor (1953)–we work with a collection of mixture spaces indexed by the alternatives. This leads us to introduce a condition that "bridges" the *gaps* in the order between the mixture spaces. The resulting representation resembles that of [vNM] in that it is both state and context independent, but is distinguished by the fact that the uniqueness of the representation depends on the number of *components* preferences generate. Methodologically speaking, the closest models in the literature to the model of this chapter is that of Karni and Safra (2000) as well as Fishburn's multilinear utility model.

Our main conclusion is that when preferences vary with context, unless the decision-maker is willing to extend the domain of preferences to include the context space, we should only hope for preferences to have a linear utility representation (as in expected utility) in mathematically interesting but behaviourally exceptional circumstances. If instead there is a willingness to forego either the linearity property or the utility property of the representation *across context space*, then the conditions we need to impose upon preferences are much more natural and capture a far wider range of reasonable behaviour.

1.A Measurability and comparability

One of the most widely used concepts in economics is that of a function that represents preferences and allows the analysis of the agent's behaviour to extend beyond the realms of set theory and make use of calculus, probability theory and many other areas of mathematics. Such a function is referred to as a utility function, and most often denoted by u. It is said to represent preferences, $\succ$, if the values it assigns to any given element of the choice set is strictly greater than the value it assigns to any other if and only if the agent strictly prefers the former to the latter. In other words, if A is the choice set, and $u : A \longrightarrow \mathcal{R}$, $a \mapsto u(a)$ if for every a, b in A:

$$a \succ b \qquad \Leftrightarrow \qquad u(a) > u(b). \tag{1.4}$$

When such a representation exists it constitutes an order embedding of $(A, \succ)$ into $(\mathcal{R}, >)$. Alternatively $(A, \succ)$ is order isomorphic to $(u(A), >)$, where $u(A)$ is the image of u in $\mathcal{R}$. It is one of the main objectives of decision theory to find conditions on preferences that are appealing in either or both the normative and positive sense, and such that they are necessary and sufficient for a representation to exist.

Closely related to the question of existence is of course the question of uniqueness. As a representation that arises from conditions on preferences will rarely be unique, it is natural to seek to identify the equivalence class to which the representation belongs. The dual to this problem is that of identifying the largest class of transformations of a utility function such that, for every transformation in the class, the transformed utility function still represents preferences. That is to say, if v is another function that represents $\succ$, then what kind of transformation relates v to u?

The largest class of transformations that preserves the representation is of course the class of strictly monotone transformations. In particular, if u and v represent the same preferences then $u = \psi \circ v$ for some strictly monotonically increasing $\psi : \mathcal{R} \longrightarrow \mathcal{R}$. Preferences that admit a representation which is unique only upto a strictly monotone transformation are said to be represented by an *ordinal utility function* from A to $\mathcal{R}$. Alternatively, preferences are said to be *ordinally measurable.*

In choice under risk or uncertainty it seems reasonable to assume that some concept of likelihood or probability (and as a result integral/summation) plays a significant role in the decision, and in this setting, as we will see, the category of ordinal functions may prove too large to be practical. Instead, it is standard practice to consider the class of *cardinal utility* functions which are defined to be those that are unique upto a strictly positive affine transformation. We say that such preferences are *cardinally measurable.*

We now define the relevant concepts for understanding the uniqueness properties of a representation for decisions under uncertainty. Given a particular degree of measurability, there are various degrees of comparability. We will now define the concepts we need to understand the following three: cardinally measurable, non-comparable; cardinally measurable, fully comparable and cardinally measurable, unit-comparable. For a detailed exposition of the these concepts in the context of social choice see Senn (1973), Roberts (1980), Blackorby et al. (1984) or d'Aspremont and Gevers (2001); indeed the definitions we introduce below are based upon Roberts (1980).

Definition 1.5. *An* invariance transformation *is a list of functions $\psi = (\psi_s, s \in S)$ with the property that for all u, $u' \in \mathcal{U}$, if for all $a \in A$, $s \in S$, $u'_s(a) = \psi_s(u_s(a))$ then u and u' represent the same family of preferences, $(\succ_s, s \in S)$. The* invariance class, *Ψ, corresponding to $(\succ_s, s \in S)$, is the set of invariance transformations.*

Definition 1.6 (CNC)**.** *Suppose the family of preferences, $(\succ_s, s \in S)$ is represented by the family of state-utilities $(u_s, s \in S)$. Preferences are said to be cardinally-measurable and non-comparable across states if $\psi \in \Psi_N$ iff it is a list of m independent, strictly positive, affine transformations. That is $\psi \in \Psi_N$ if and only if:*

$$\forall (a, s) \in A \times S, \quad \psi_s(u_s(a)) := k_s + l_s u_s(a) \textit{ for some } k_s \in \mathcal{R} \textit{ and } l_s > 0.$$

Remark 1.2. *In the theory of measurement, of which the theory of utility or welfare measurement is a special case, the terms $k_.$ and $l_.$ of the affine transformation have a specific meaning. The "intercept" $k_.$ is referred to as the origin of the scale, whilst the "slope" $l_.$ is referred to as the unit of the scale or as the unit of measurement.*

Remark 1.3. *Note that in the definition of (CNC), for each state s, both the unit l_s and the origin k_s of the scale is independent of the unit l_t and the origin k_t of the scale of every other state t. This means that if preferences are measurable up to this degree, then the only kind of comparison of welfare/utility across elements of Y that has meaning is the kind where we fix the state s and take the ratio of one difference of utilities of an act-pair to another.*

To see this, suppose u is a family of state-utility functions representing the family of preferences $(\succ_s, s \in S)$, and that preferences are CNC. For $r = s, t \in S$, suppose we consider the equivalence class of representations of the preferences

$$\psi_r \circ u_r := k_r + l_r u_r, \text{ for each } k_r \in \mathcal{R} \text{ and } 0 < l_r < \infty.$$

Then provided $c \not\sim_t d$ as ψ varies across the invariance class Ψ_N, the equality in the following derivation holds if and only if $l_s = l_t$:

$$\begin{aligned}
\frac{\psi_s \circ u_s(a) - \psi_s \circ u_s(b)}{\psi_t \circ u_t(c) - \psi_t \circ u_t(d)} &\equiv \frac{\big((k_s + l_s u_s(a)) - (k_s + l_s u_s(b))\big)}{\big((k_t + l_t u_t(c)) - (k_t + l_t u_s(d))\big)} \\
&\equiv \frac{l_s\big(u_s(a) - u_s(b)\big)}{l_t\big(u_t(c) - u_t(d)\big)} \\
&= \frac{u_s(a) - u_s(b)}{u_t(c) - u_t(d)},
\end{aligned}$$

so that it is infact an identity over Ψ_N if and only if $s = t$, and in this case the ratio is equal to some unique real number z. If z is greater than 1 we know that in state s the utility of a less the utility of b is greater than the utility of c less the utility of d. If $s \neq t$, then precisely because the above ratio is not constant, we can make no such comparison, even when $c = a$ and $d = b$.

A widely accepted scenario where a single decision-maker's preferences may be incomparable across states is when the uncertainty concerns states of health (Karni 1985 p45, MWG p199). In the literature it is referred to as state-dependent preferences, this is somewhat misleading as preferences over the set of alternatives may well be state-dependent yet comparable. We will return to this point in the next section.

Definition 1.7 (CFC)**.** *Suppose the family of preferences,* $(\succ_s, s \in S)$ *is represented by the family of state-utilities* $(u_s, s \in S)$*. Preferences are said to be cardinally-measurable and fully-comparable across states if* $\psi \in \Psi_F$ *iff it is a list of* m *identical, strictly positive, affine transformations. That is* $\psi \in \Psi_F$ *iff:*

$$\forall (a, s) \in A \times S, \quad \psi_s(u_s(a)) := k + l u_s(a) \text{ for some } k \in \mathcal{R} \text{ and } l > 0.$$

Remark 1.4. *For this invariance class, comparisons of ratios of differences in utilities between arbitrary pairs of elements of the space* $A \times S$ *are possible.*

Between these two extreme forms of cardinal measurability, we have the following definition.

Definition 1.8 (CUC)**.** *Suppose the family of preferences,* $(\succ_s, s \in S)$ *is represented by the family of state-utilities* $(u_s, s \in S)$*. Preferences are said to be cardinally-measurable and unit-comparable across states if* $\psi \in \Psi_U$ *iff it is a list of* m *strictly positive, affine transformations such that across states the unit of measurement is identical and the origin is independent. That is* $\psi \in \Psi_U$ *iff:*

$$\forall (a, s) \in A \times S, \quad \psi_s(u_s(a)) := k_s + l u_s(a) \text{ for some } k_s \in \mathcal{R} \text{ and } l > 0,$$

1.B Embedding Δ in $\mathcal{R}^n$ and $\mathcal{R}^{n-1}$

The next two results show that given our space of contexts is the set of probability distributions over the finite state-space S, we may embed Δ in $\mathcal{R}^{n-1}$. Their statement and proof is provided for the sake of clarity and completeness of the argument.

Theorem 1.2 (Evans (2010)). *Suppose that K is a non-empty, metrizable, compact, convex subset of a locally convex space E. The following assertions are equivalent*

(1) *K is a simplex.*

(2) *For all a' , $a'' \geqslant 0$ and $b', b'' \in E$, either $(a'K + b') \cap (a'K + b')$ is empty, or of the form $aK + b$ for some $a \geqslant 0$ and $b \in E$.*

(3) *For each $x \in K$ there is a unique probability measure μ supported on $\operatorname{ext} K$ such that μ has barycenter x.*

Proposition 1.1. *Let*

$$\Delta' := \left\{x \in \mathcal{R}^n_+ : \sum_{s=1}^{n} x_s = 1\right\} \quad \textit{and} \quad \Delta'' := \left\{x \in \mathcal{R}^{n-1}_+ : \sum_{s=1}^{n-1} x_s \leqslant 1\right\}.$$

The set Δ of probability distributions over the set $S \equiv \{1, \ldots, n\}$ is homeomorphic to both of these sets. Moreover, $\operatorname{aff}(\Delta')$ and $\operatorname{aff}(\Delta'')$ are homeomorphic under the affine transformation

$$e_n : \mathcal{R}^{n-1} \longrightarrow \mathcal{R}^n, \quad x \longmapsto i(x) + \left(1 - \sum_{s=1}^{n-1} x_s\right)\delta'_n$$

where $i : \mathcal{R}^{n-1} \longrightarrow \mathcal{R}^{n-1} \times \{0\} \subset \mathcal{R}^n$, $x = (x_1, \ldots, x_{n-1})^* \mapsto (x_1, \ldots, x_{n-1}, 0)^*$

and $\delta'_n = (0, \ldots, 0, 1)^*$ *is the element of the standard basis for* $\mathcal{R}^n$ *corresponding to the* n^{th} *dimension.*[14]

Proof. We state without proof that both Δ' and Δ'' are convex, $n-1$-dimensional with extremal points $\{\delta'_s\}_{s\in S}$ and $\{\delta''_s\}_{s\neq n} \cup \{0\}$ respectively, and that the cardinality of each of the latter two sets n. Thus by theorem (1.2), we know that for each $p' \in \Delta'$, there exists a unique probability distribution $p = \{p_1, \ldots, p_n\} \in \Delta$, over S such that

$$p' = \sum_{s\in S} p_s \delta'_s.$$

Similarly, for each $p'' \in \Delta''$, there exists a unique probability distribution $q = \{q_1, \ldots, q_n\} \in \Delta$ over S such that

$$p'' = \sum_{s\in S\backslash n} q_s \delta''_s + q_n 0,$$

Therefore, to each of the sets Δ' and Δ'', there exists a (continuous) bijective function (with continuous inverse) from Δ.

From the above, we also conclude that $\Delta' = \text{conv}(\delta'_1, \ldots, \delta'_n)$ and $\Delta'' = \text{conv}(\delta''_1, \ldots, \delta''_{n-1}), 0)$. Similarly, $\text{aff}(\Delta') = \text{aff}(\delta'_1, \ldots, \delta'_n)$, and $\{\delta''_s\}_{s\neq n} \cup \{0\}$ forms a basis for $\text{aff}(\Delta'') = \mathcal{R}^{n-1}$. We now prove that e_n is indeed an affine transformation which is a bijection from $\mathcal{R}^{n-1}$ to $\text{aff}(\Delta')$. This will complete the proof since an affine transformation is continuous.

[14] Note that x^* denotes the transpose of the vector x, and we will denote by δ'_s the element of the standard basis for $\mathcal{R}^n$ corresponding to the s^{th} dimension, for $s \in S$. We similarly define δ''_s as a basis vector in $\mathcal{R}^{n-1}$.

The fact that e_n is injective follows immediately from the fact that if $x, y \in \mathcal{R}^{n-1}$, $x \neq y$, then there exists $s \in \{1, \ldots, n-1\}$ such that $x_s \neq y_s$. That is, because the definition of i implies that $i(x)_s \neq i(y)_s$ for this value of s, and because δ_n is has a zero entries for all but the n^{th} element, $e_n(x)$ is not equal to $e_n(y)$.

To see that it is surjective, take $x' \in \text{aff}(\Delta')$. Then because $\text{aff}(\Delta') = \text{aff}(\delta_1, \ldots, \delta_n)$, we have $x'_1, \ldots, x'_n = 1$ and so

$$\begin{aligned} x' &= x'_1\delta'_1 + \cdots + x'_n\delta'_n \\ &= x'_1\delta'_1 + \cdots + x'_{n-1}\delta'_{n-1} + \big(1 - \sum_{s=1}^{n-1} x'_s\big)\delta'_n. \end{aligned}$$

Now since $\sum_{s=1}^{n-1} x'_s\delta'_s = (x'_1 \ldots, x'_{n-1}, 0)^*$, we know that this element lies in the image of i. That is, $(x'_1 \ldots, x'_{n-1})$ lies in $\mathcal{R}^{n-1}$, and moreover, its image under e_n is equal to x'.

We now show that e_n is indeed affine. For any $x = (x_1, \ldots, x_{n-1})^* \in \mathcal{R}^{n-1}$ we define x_n to be $1 - x_1 - \cdots - x_{n-1}$. Then for all $\alpha \in \mathcal{R}$, and $y, z \in \mathcal{R}^{n-1}$, by the linearity of i and the definition of y_n and z_n we have

$$\begin{aligned} \alpha e_n(y) + (1-\alpha)e_n(z) &= \alpha i(\alpha y + (1-\alpha)z) + \big(\alpha y_n(1-\alpha)z_n)\big)\delta_n \\ &= \alpha i(\alpha y + (1-\alpha)z) + \big(1 - \sum_{s=1}^{n-1}(\alpha y_s + (1-\alpha)z_s)\big)\delta_n \\ &= e_n(\alpha y + (1-\alpha)z), \end{aligned}$$

since $\alpha y + (1-\alpha)z \in \mathcal{R}^{n-1}$ and the sum of the elements of this vector and $1 - \sum_{s=1}^{n-1}(\alpha y_s + (1-\alpha)z_s)$ is one. Thus, as required for the definition of affine transformation, e_n takes points on a line to points on a line and it preserves midpoints (simply take $\alpha = \frac{1}{2}$). □

Corollary 1.1. *Each probability distribution p in Δ uniquely identifies elements $p' \in \Delta'$ and $p'' \in \Delta''$ such that $e_n(p'') = p'$.*

Proof. This follows immediately from definition of e_n, the above proposition, and the fact that probability distributions sum to one, so that the weight p_n assigned to state n is $1 - \sum_{s=1}^{n-1} p_s$. □

Remark 1.5. *A consequence of the two results is that we may represent preferences $\{A, \succ_p \mid p \in \Delta\}$ by either $\{A, \succ'_{p'} \mid p' \in \Delta'\}$ or $\{A, \succ''_{p''} \mid p'' \in \Delta''\}$, where $\succ'_{\cdot}$ and $\succ''_{\cdot}$ satisfy: for each $p \in \Delta$, $p' \in \Delta'$ and $p'' \in \Delta''$ the following is true for all $a, b \in \Delta$*

$$a \succ_p b \quad \Leftrightarrow \quad a \succ'_{p'} b \quad \Leftrightarrow \quad a \succ''_{p''} b,$$

whenever $p = \{p_1, \ldots, p_n\}$, $p' = \sum_{s=1}^{n} p_s\, \delta'_s$, and $p'' = \sum_{s=1}^{n-1} p_s\, \delta''_s$.

Henceforth whenever there is no possibility of confusion we will refer to either Δ' or Δ'', simply as Δ.

Chapter 2

A geometric approach to representing context preferences

2.1 Motivation

The model we present in this chapter is closely related to Gilboa and Schmeidler (2003) ([GS03]). In that paper the authors provide a re-interpretation of their model of *case-based decision* theory (Gilboa and Schmeidler (1995, 2002) and work in the same context space as we do in this thesis. That is, where states play the role of cases and probabilities or beliefs play the role of context. Whilst the results of this chapter only apply to the context space that is the set of probability distributions over a finite state-space, the proofs of Gilboa and Schmeidler (1995), (2001) and (2003) suggest that there is no reason why they cannot be extended to the type of context space we find in case-based decision theory.

The main reason why we propose the present model as an alternative to [vNM] is the following. When the task of ranking alternatives is nontrivial, the model that defines preferences over the smallest possible domain, whilst still giving rise to an appropriate expected utility representation, should, all other things being equal, dominate other models. Moreover, expected utility representations are most appealing for decision problems where an agent anticipates being in a position of uncertainty about the future, is prepared to define an order over the set of alternatives for each of the anticipated positions of uncertainty, and does so in a consistent way. The model of [vNM], by defining preferences on mixtures of outcomes, in general demands a good deal more information regarding preferences than this.

We argue that especially when the uncertainty concerns the actions of a non-strategic actor such as nature, this excess information is of no use to the decision maker. The model proposed here defines preferences on the smallest possible domain for preferences that will guarantee the existence of a single expected utility function that faithfully represents preferences conditional upon any of the anticipated positions of uncertainty. The cost of doing so is that preferences must, in an appropriate sense, be diverse.

A different and equally appealing justification for the model in this chapter is provided by [GS03]. There the authors provide a convincing argument of why it may be implausible to define preferences over lotteries on observable consequences in certain types of game both against nature and opponents. The games the authors have in mind are the dictator and ultimatum games where players are often observed playing dominated strategies. In the same spirit as Hammond (1989), they suggest that "extended" or psychological consequences, which cannot be explicitly simulated in experiments may well be giving rise to the apparent failure of expected utility to model behaviour. In contrast with [vNM], they argue for a model of expected utility "in the context of the game".

This chapter improves upon the model of [GS03] in two respects. Firstly, we weaken the diversity condition in an important way. The resulting diversity condition has two parts the first of which has been anticipated by Ashkenazi and Lehrer (2001). Together the two diversity conditions are somewhat weaker than that of [GS03], allowing for diagrammatic representation of a variety of preferences that were previously excluded when for instance the number of states is equal to 3. In [GS03] p.189 the authors state that they "do not know of a set of axioms that are necessary and sufficient for a representation ... by a matrix u that need not be diversified". This chapter provides just such a set of conditions on preferences. Moreover the conditions make precise the two distinct issues that arise when we seek a linear utility representation of preferences when the context space is the positive orthant. Secondly, we provide an alternative to the "combination" condition of [GS03] that gives rise to the linear form of the representation. The resulting conditions are weaker and we argue more intuitive. The main theorem of this chapter is, like that of [GS03], a statement of equivalence between conditions on preferences and an expected utility representation with specific properties.

The exposition consists of three parts. The first is where the implications of imposing conditions on pairs of alternatives are considered and the second is where conditions triples and quadruples are considered and where the main representation theorem is presented. Finally, in the third part a detailed comparison with the model of [GS03] is made.

2.2 Conditions on pairs of alternatives across contexts

We will now look at the extent to which it is possible to represent context preferences $(A, \succ_p)_{p \in \Delta}$ when the following conditions are imposed.

Definition (Asymmetry (Asy.)).
For all $a, b \in A$, $p \in \Delta$: *if* $a \succ_p b$ *then* $\neg(b \succ_p a)$; *equivalently,*

$$p \in \mathcal{B}_{ab} \Rightarrow p \notin \mathcal{W}_{ab}.$$

Definition (Weak Pareto dominance (WP)).
For all p *in* Δ : *if* $\neg(a \succ_s b)$ *for all* $s \in \operatorname{supp}(p)$, *then* $\neg(a \succ_p b)$; *equivalently,*

$$p \in \mathcal{B}_{ab} \ \Rightarrow \ \exists s \in \operatorname{supp}(p) : \delta_s \in \mathcal{B}_{ab}.$$

Definition (Continuity (C'ty)).
For all $a, b \in A$, $p \in \Delta$: *if* $a \succ_p b$, *then there exists an open neighbourhood* U *of* p *in* Δ *such that for every* $q \in U$ *we have* $a \succ_q b$.

Definition (Strong betweenness across contexts (p-SB)).
For all $a, b \in A$, $p, q \in \Delta$, *and* $0 < \lambda < 1$, *let* $r := \lambda p + (1 - \lambda) q$. *If* $\neg(b \succ_p a)$ *and* $a \succ_q b$, *then* $a \succ_r b$; *equivalently,*

$$p \in \Delta \backslash \mathcal{W}_{ab} \ \wedge \ q \in \mathcal{B}_{ab} \Rightarrow r \in \mathcal{B}_{ab}.$$

Taken together, (p-SB) and (WP) are a weakening of the combination condition of [GS03]. This condition is written in the following way, where $\gtrsim_p = \succ_p \cup \sim_p$. (It is clear that by definition of $\sim_p$ this is in fact a disjoint union, which henceforth will be written as $\dot\cup$.)

Definition 2.1 (Combination (Comb.) (Gilboa and Schmeidler (2003))**.** *For all $a, b \in A$, $p, q \in \Delta$, and $0 < \lambda < 1$, let $r := \lambda p + (1-\lambda)q$. If $a \gtrsim_p b$ $(a \succ_p b)$ and $a \gtrsim_q b$, then $a \gtrsim_r b$ $(a \succ_r b)$; equivalently, both*

$$p \in \Delta \backslash \mathcal{W}_{ab} \ \ \wedge \ \ q \in \Delta \backslash \mathcal{W}_{ab} \Rightarrow r \in \Delta \backslash \mathcal{W}_{ab},$$

and

$$p \in \mathcal{B}_{ab} \ \ \wedge \ \ q \in \Delta \backslash \mathcal{B}_{ab} \Rightarrow r \in \mathcal{B}_{ab}.$$

It is clear that the second of the two statements in condition (Comb.) is equivalent to (p-SB) and the first is stronger than (WP). (Although we omit a proof of this claim, the argument follows by an argument which we present in the next chapter: proposition (3.2). Indeed, it is clear that condition (p-SB) is more than just combination or betweenness as described in Chew (1989) for instance. It contains what may intuitively be described as a thinness property of $\mathcal{N}_{ab}$ for each $a, b \in A$, the discussion of which we defer to chapter 3.

We now provide the definition of three kinds of separation by a hyperplane that are used in the sequel.

Definition 2.2 (Webster (1994))**.** *Let H, A and B be subsets of $\mathcal{R}^n$ with H being a hyperplane.*

- *H is said to separate A and B if A lies in one of the closed half-spaces* *determined by H, and B lies in the other.*

- *H is said to properly separate A and B if it separates them but does not contain them both.*

- *H is said to strictly separate A and B if A lies in one of the open half-spaces determined by H and B lies in the other.*

The next result is the main mathematical result of this section, and the representation theorem of this section follows as a corollary.

Proposition 2.1. *Let conditions (Asy.), (WP), (C'ty) and (p-SB) hold for the familly of ordered sets $(A, \succ_p)_{p \in \Delta}$. Then for all $a, b \in A$, if $\mathcal{B}_{ab}$ is nonempty, then there exists a hyperplane H in $\mathcal{R}^{n-1}$ such that*

$$H \cap \Delta = \mathcal{N}_{ab}$$

Furthermore, when both $\mathcal{B}_{ab}$ and $\mathcal{W}_{ab}$ are nonempty, H is the unique hyperplane that supports each set and strictly separates them.

Proof. Throughout the proof, we consider an arbitrary pair of elements $a, b \in A$. We first consider the case where $\mathcal{W}_{ab}$ is empty. The proof for this case is presented as follows: step (i) we show that $\mathcal{N}_{ab} \subset H$ for some hyperplane in $\mathcal{R}^{n-1}$, and in step (ii) we show that $H \cap \Delta \subset \mathcal{N}_{ab}$. For the case where $\mathcal{W}_{ab}$ is nonempty, once again (i) we first show that $\mathcal{N}_{ab} \subset H$ for some hyperplane in $\mathcal{R}^{n-1}$ that supports $\mathcal{B}_{ab}$ and $\mathcal{W}_{ab}$, (ii) we then show that H properly separates these sets, and then (iii) we show that $H \cap \Delta \subset \mathcal{N}_{ab}$.

Case ($\mathcal{W}_{ab} = \varnothing$), step (i): Note that if $\mathcal{N}_{ab}$ is empty, then the proposition is satisfied for any hyperplane H that does not intersect Δ. Thus suppose $\mathcal{N}_{ab}$ is nonempty. Then since $\mathcal{N}_{ab} \neq \Delta$, $\Delta = \mathcal{N}_{ab} \dot{\cup} \mathcal{B}_{ab}$, where $\mathcal{B}_{ab}$ is nonempty. If we let $T := \{s \in S : a \sim_p b\}$, then condition (WP) implies that $\Delta_T := \mathrm{conv}(\{\delta_s : s \in T\})$ is a subset of $\mathcal{N}_{ab}$.

We now show that condition (p-SB) implies that Δ_T contains $\mathcal{N}_{ab}$. Suppose $p \in \mathcal{N}_{ab} \backslash \Delta_T \neq \varnothing$. Note that the set T is maximal in the sense that it contains all the elements s of S such that $a \sim_s b$ holds. Moreover, as $\mathcal{B}_{ab}$ is nonempty, condition (WP) implies that $S \backslash T$ is nonempty, so we know that every extremal measure in $\Delta \backslash \Delta_T$ is contained in $\mathcal{B}_{ab}$. This implies that because p lies outside Δ_T, it cannot be extremal. Furthermore, because Δ_T is closed and convex, it is equal to its convex hull, thus p cannot be written as a convex combination of the extremal elements of Δ_T. Condition (p-SB) implies that $\Delta_{S \backslash T}$ is a subset of $\mathcal{B}_{ab}$, so p lies outside the set of probability measures with support $S \backslash T$. Therefore, $\mathrm{supp}(p)$ contains elements of both T and its complement in S.

Therefore without loss of generality we may suppose

$$p = \sum_{s=1}^{n} \lambda_s \delta_s,$$

where for $s = 1, \ldots k$, $s \in T$ and for $s = k+1, \ldots n$, $s \in S \backslash T$, and the λ_s sum to one. (To mitigate the fact that we are in $\mathcal{R}^{n-1}$, we assume, without loss of generality, that $\delta_n = 0 \in \mathcal{R}^{n-1}$, and let $\lambda_n \equiv 1 - \sum_{s=1}^{n-1} \lambda_s$.)

Let $\lambda' := \sum_{s=1}^{k} \lambda_s$ and $\lambda'' := \sum_{s=k+1}^{n} \lambda_s$. By the argument of the preceding paragraph, neither λ' nor λ'' equals zero. If we let $q' := \frac{1}{\lambda'}\sum_{s=1}^{k}\lambda_s\delta_s$ and $q'' := \frac{1}{\lambda''}\sum_{s=k+1}^{n}\lambda_s\delta_s$, then q' and q'' clearly lie in Δ_T and $\Delta_{S\setminus T}$ respectively. Now consider the affine hull of q' and p in $\mathcal{R}^{n-1}$. This may be represented by $r : \mathcal{R} \longrightarrow \mathcal{R}^{n-1}$, $\nu \mapsto r(\nu)$ where

$$r(\nu) := \nu p + (1-\nu)q',$$

and so by the definition of p and q', we have

$$\begin{aligned} r(\nu) &= \nu\Big(\sum_{s=1}^{k}\lambda_s\delta_s + \sum_{s=k+1}^{n}\lambda_s\delta_s\Big) + (1-\nu)\frac{1}{\lambda'}\sum_{s=1}^{k}\lambda_s\delta_s \\ &= \nu\sum_{s=k+1}^{n}\lambda_s\delta_s + \Big((1-\nu)\frac{1}{\lambda'} + \nu\Big)\sum_{s=1}^{k}\lambda_s\delta_s. \end{aligned}$$

Now if $(1-\nu)\frac{1}{\lambda'} + \nu = 0$ whenever $\nu = 1/\lambda''$, then we have

$$r(\nu) = q''.$$

The following computation shows that this is indeed true because $\lambda' + \lambda'' = 1$:

$$\begin{aligned} (1-\nu)\frac{1}{\lambda'} + \nu &= \frac{(\lambda''-1)}{\lambda''}\frac{1}{\lambda'} + \frac{1}{\lambda''} \\ &= \frac{(\lambda''-1+\lambda')}{\lambda''\lambda'}. \end{aligned}$$

Finally, the fact that $1/\lambda''$ is greater than 1 implies that $q'' = r(\frac{1}{\lambda''})$ lies further from $q' = r(0)$ in $\mathrm{aff}(p, q')$ than $p = r(1)$, so that condition (p-SB) implies that p lies in $\mathcal{B}_{ab}$, contradicting the assumption that it was an element of $\mathcal{N}_{ab}$. Therefore we conclude that $\mathcal{N}_{ab}$ is a subset of Δ_T.

Now the fact that Δ_T is contained in $H \cap \Delta$ for some hyperplane in $\mathcal{R}^{n-1}$ follows from the fact that Δ_T is a convex set together with the fact that $\dim(\Delta_T) < \dim(\Delta) = n - 1$, for then Δ_T is equal to its boundary in $\mathcal{R}^{n-1}$. Theorem 2.4.12 of Webster p.71 then implies the desired result.

Case ($\mathcal{W}_{ab} = \varnothing$) Step (ii): To identify a hyperplane in Δ that satisfies $H \cap \Delta = \mathcal{N}_{ab}$, we we first note that because $S \backslash T$ is nonempty, $k := |T| \leqslant n-1$. So the dimension of Δ_T is less than or equal to $n-2$. Note that the extremal points δ_s, $s \in T$, define a basis for $\text{aff}(\Delta_T)$, and therefore this too is of dimension $k-1$. This in turn implies that $\text{aff}(\Delta_T)$ contains no elements of Δ other than Δ_T, for if it did, then any such a point would lie outside the span of $\{\delta_s\}_{s \in S}$, and the dimension of $\text{aff}(\Delta_T)$ would necessarily be greater than $k-1$.

If $k = n-1$ then $\text{aff}(\Delta_T)$ is itself the uniquely identified hyperplane H in $\mathcal{R}^{n-1}$ that satisfies $H \cap \Delta = \mathcal{N}_{ab}$. If not then for each $s = k+1, \ldots n$, consider the affine hull of the pair $\widehat{p} := \frac{1}{k}(\delta_1, \ldots, \delta_k)$, and $\widehat{p} + \delta_s$. That is the translation by $\widehat{p}$ of the straight line in $\mathcal{R}^{n-1}$ that is spanned by δ_s for each s in $S \backslash T$. Recalling that $\delta_n = 0$, so that the affine hull of $\widehat{p}$ and $\widehat{p} + \delta_n$ is just the point $\widehat{p}$, we note that for each $k+1 \leqslant s \leqslant n-1$, the image of the function $r_s : \mathcal{R} \longrightarrow \mathcal{R}^{n-1}$, $\nu \mapsto r_s(\nu)$ is equal to $\text{aff}(\widehat{p}, \widehat{p} + \delta_s)$ whenever

$$\begin{aligned} r_s(\nu) &:= \nu(\widehat{p} + \delta_s) + (1 - \nu)\widehat{p} \\ &= \nu\delta_s + \widehat{p}. \end{aligned}$$

The definition of $\widehat{p}$ is such that for every s, $r_s(\nu)$ is a probability distribution and hence an element of Δ, if and only if $\nu = 0$, and at this value

$r_s(0) = \hat{p} \in \Delta_T$. An identical argument is true of affine combinations of elements of the set $\Delta_T \cup \{r_{k+1}(1), \ldots, r_{n-1}(1)\}$ as only affine combinations that assign zero weight to elements outside Δ_T are probability distributions. Finally, since the vectors δ_s, $s = 1, \ldots, k$ and $r_s(1)$, $s = k+1, \ldots n-1$ are a set of $n-1$ linearly independent vectors, they define the basis for the desired hyperplane H in $\mathcal{R}^{n-1}$.

Case ($\mathcal{W}_{ab} \neq \emptyset$), step (i): We know that by condition (Asy.), $\mathcal{B}_{ab}$ and $\mathcal{W}_{ab}$ are disjoint, and by condition (C'ty) open subsets of the connected set Δ. Moreover, because of the fact that $\mathcal{W}_{ab} \equiv \mathcal{B}_{ba}$, we only need to consider the sub-case where $\mathcal{B}_{ab}$ is also nonempty, so that by proposition (3.1.20) they are topologically separated sets, that are topologically separated by

$$\mathcal{N}_{ab} = \Delta \backslash (\mathcal{B}_{ab} \cup \mathcal{W}_{ab})$$

which is therefore also a nonempty set.

We will first show that $\mathcal{N}_{ab}$ is a subset of both $\mathrm{cl}_\Delta(\mathcal{B}_{ab})$ and $\mathrm{cl}_\Delta(\mathcal{W}_{ab})$. Then, because condition (p-SB) implies that $\mathcal{B}_{ab}$ is convex, theorem (2.3.5) of Webster p.62 states that $\mathrm{cl}_\Delta(\mathcal{B}_{ab})$ is also convex. Moreover, the symmetry between $\mathcal{B}_{ab}$ and $\mathcal{W}_{ab}$ ($\mathcal{W}_{ab} \equiv \mathcal{B}_{ba}$), will allow us to conclude that

$$\mathcal{N}_{ab} \subset \mathrm{cl}_\Delta(\mathcal{B}_{ab}) \cap \mathrm{cl}_\Delta(\mathcal{W}_{ab})$$

Then by theorem (2.1.3) of Webster p.50, we know that arbitrary intersections of convex sets are convex. The fact that $\mathcal{N}_{ab}$ is equal to this intersection will then follow immediately from the fact $\mathcal{B}_{ab}$ and $\mathcal{W}_{ab}$ are topologically separated, and so neither contains closure points of the other. This will allow us to conclude that $\mathcal{N}_{ab}$ is convex.

We now show that $\mathcal{N}_{ab} \subset \mathrm{cl}_\Delta(\mathcal{B}_{ab})$. Take any element p in $\mathcal{N}_{ab}$ and q in $\mathcal{B}_{ab}$. By condition (p-SB), we know that $q\frac{1}{2}p \equiv \frac{1}{2}q + \frac{1}{2}p$ is also a element of $\mathcal{B}_{ab}$. Then, by the same condition, the same is also true of $q\frac{1}{4}p \equiv \frac{1}{4}q + \frac{3}{4}p$. Indeed, it is true for the entire sequence $q(\frac{1}{2})^j p$, $j \in \mathbb{N}$. Moreover, since this sequence converges (in the weak* topology on Δ) to p, the latter is an element the closure of $\mathcal{B}_{ab}$.

This argument not only proves that $\mathcal{N}_{ab}$ is a subset of $\mathrm{cl}_\Delta \mathcal{B}_{ab}$, but also, since condition (C'ty) implies that $\mathcal{B}_{ab}$ is open in Δ, $\mathcal{N}_{ab}$ is equal to the boundary of $\mathcal{B}_{ab}$ in Δ, $\mathrm{bd}_\Delta(\mathcal{B}_{ab})$. By symmetry, the same is true of $\mathcal{W}_{ab} = \Delta \backslash \mathrm{cl}_\Delta(\mathcal{B}_{ab})$. Therefore, $\mathcal{N}_{ab}$ is equal to the common boundary of $\mathcal{B}_{ab}$ and $\mathcal{W}_{ab}$ in Δ. Now since $\mathcal{B}_{ab}$ has nonempty interior relative to $\mathcal{R}^{n-1}$, and Δ is a closed subset of $\mathcal{R}^{n-1}$ lemma (3.1.32) applied with $Z = \mathcal{R}^{n-1}$, $Y = \Delta$ and $\mathcal{B}_{ab} = X$ implies that the boundary of $\mathcal{B}_{ab}$ relative to $\mathcal{R}^{n-1}$ contains the boundary of $\mathcal{B}_{ab}$ relative to Δ, and the same is true for $\mathcal{W}_{ab}$.

We have therefore shown that $\mathcal{N}_{ab}$ is a convex subset of

$$\mathrm{rebd}(\mathcal{B}_{ab}) \cap \mathrm{rebd}(\mathcal{W}_{ab}),$$

and is equal to its own boundary in $\mathcal{R}^{n-1}$. Theorem (2.4.12) of Webster p.71 then implies that there exists a hyperplane H in $\mathcal{R}^{n-1}$ that contains $\mathcal{N}_{ab}$, and nontrivially supports both $\mathcal{B}_{ab}$ and $\mathcal{W}_{ab}$.

Case ($\mathcal{W}_{ab} \neq \emptyset$), step (ii): To see that H properly separates $\mathcal{B}_{ab}$ and $\mathcal{W}_{ab}$, we first show that $\mathcal{N}_{ab}$ (and hence H) has nonempty intersection with $\mathrm{ri}\,\Delta$. Suppose $\mathcal{N}_{ab} \subset \mathrm{rebd}(\Delta)$. Then since $\mathrm{ri}\,\Delta \backslash \mathcal{N}_{ab} = \mathrm{ri}\,\Delta$, $\mathcal{N}_{ab}$ does not topologically separate the connected set $\mathrm{ri}\,\Delta$. Now note that we have

$$\mathrm{ri}\,\Delta \subset \Delta \backslash \mathcal{N}_{ab} \subset \Delta,$$

so that, by theorem (23.4) of Munkres p.166, $\Delta \backslash \mathcal{N}_{ab} = \mathcal{B}_{ab} \cup \mathcal{W}_{ab}$ is a connected set. This of course contradicts conditions (Asy.) and (C'ty) which together imply that $\mathcal{B}_{ab}$ and $\mathcal{W}_{ab}$ are disjoint open sets.

Suppose that H does not separate $\mathcal{B}_{ab}$ and $\mathcal{W}_{ab}$. Then because H supports each set, $\mathcal{B}_{ab} \cup \mathcal{W}_{ab}$ is contained in only one of the half spaces determined by H. However the fact that $\mathcal{N}_{ab}$ is a subset of H implies that Δ is contained in only one of the closed half spaces determined by H. But H has nonempty intersection with $\mathrm{ri}(\Delta)$, and as $\mathrm{ri}(\Delta)$ is open in $\mathcal{R}^{n-1}$ it intersects both the half-spaces determined by H. This is the desired contradiction, and thus H separates $\mathcal{B}_{ab}$ and $\mathcal{W}_{ab}$. Indeed, because each of these sets is open, neither lies in H, and so H properly separates $\mathcal{B}_{ab}$ and $\mathcal{W}_{ab}$.

Case ($\mathcal{W}_{ab} \neq \emptyset$), step (iii): To show that $\Delta \cap H$ is a subset of $\mathcal{N}_{ab}$, the first step is to show that $\mathrm{ri}\,\Delta \cap H$ is a subset of $\mathcal{N}_{ab}$. The remainder of the proof follows by noting that since $\mathrm{ri}\,\Delta \cap H$ is nonempty and Δ is closed, $\Delta \cap H = \mathrm{cl}_{\Delta \cap H}(\mathrm{ri}\,\Delta \cap H)$, and the desired conclusion then follows by noting that for arbitrary sets $X \subset Y \subset Z$, we have $\mathrm{cl}_Y X = Y \cap \mathrm{cl}_Z X$, so that

$$\mathrm{cl}_{\Delta \cap H}(\mathrm{ri}\,\Delta \cap H) = (\Delta \cap H) \cap \mathrm{cl}_{\Delta}(\mathrm{ri}\,\Delta \cap H),$$

where $\mathrm{cl}_{\Delta}(\mathrm{ri}\,\Delta \cap H)$ is a subset of $\mathcal{N}_{ab}$ as the latter is closed in Δ.

To see that $\mathrm{ri}\,\Delta \cap H$ is indeed a subset of $\mathcal{N}_{ab}$, suppose otherwise and obtain a contradiction. Take $p \in (\mathrm{ri}\,\Delta \cap H) \cap \mathcal{B}_{ab}$. Note that since $\mathcal{B}_{ab}$ is open in Δ, it is open in any subspace (such as $\mathrm{ri}\,\Delta$) of Δ. Then since $\mathrm{ri}\,\Delta$ is open in $\mathcal{R}^{n-1}$, any open subset of $\mathrm{ri}\,\Delta$ is open in $\mathcal{R}^{n-1}$. Thus p is an element of $\mathrm{ri}\,\Delta \cap \mathcal{B}_{ab}$, and this set is open in $\mathcal{R}^{n-1}$. Since p lies in H, every open set in $\mathcal{R}^{n-1}$ containing p intersects both the open half-spaces generated by H. This contradicts the fact that H supports $\mathcal{B}_{ab}$. In the same way, p cannot lie in $\mathcal{W}_{ab}$, for recall that $\mathcal{B}_{ba} \equiv \mathcal{W}_{ab}$ and that a and b were arbitrary.

For uniqueness of H, we recall that $H \cap \mathrm{ri}\,\Delta$ is nonempty, and because $\mathrm{ri}\,\Delta$ is open in $\mathcal{R}^{n-1}$, $H \cap \mathrm{ri}\,\Delta$ is open in H.[1] Therefore, $\mathrm{aff}(H \cap \mathrm{ri}\,\Delta) = \mathrm{aff}(\mathcal{N}_{ab})$ is equal to H which is unique because $\mathcal{N}_{ab}$ is unique. □

Definition 2.3. *For every $a, b \in A$, define $\mathcal{H}_{ab}$ to be a hyperplane in $\mathcal{R}^{n-1}$ satisfying $\mathcal{H}_{ab} \cap \Delta_0 = \mathcal{N}_{ab}$. Now, using the affine homeomorphism e_n defined in proposition (1.1), let $\mathcal{H}'_{ab} := e_n(\mathcal{H}_{ab})$ in $\mathrm{aff}\,\Delta \subset \mathcal{R}^n$. Finally, define $\mathcal{I}_{ab}$ to be the affine hull of $\mathcal{H}'_{ab} \cup \{0\}$ in $\mathcal{R}^n$.*

Lemma 2.1. *For all $a, b \in A$, given a choice of $\mathcal{H}_{ab}$ in $\mathcal{R}^{n-1}$, defined as in definition (2.3), $\mathcal{I}_{ab}$ is a uniquely identified $(n-1)$-dimensional linear subspace*

[1] Then there exists a homeomorphism between $H \cap \mathrm{ri}\,\Delta$ and H, and because dimension is a topological invariant of Euclidean space (see Hurewicz and Wallman (1948)), $H \cap \mathrm{ri}\,\Delta$ is of the same dimension as H.

of $\mathcal{R}^n$. Moreover, $\mathcal{B}_{ab}$ and $\mathcal{W}_{ab}$ are strictly separated by $\mathcal{I}_{ab}$ and $\mathcal{I}_{ab} \cap \Delta = \mathcal{N}_{ab}$.

Proof. For the first statement, note that given a choice of $\mathcal{H}_{ab}$, $\mathcal{H}'_{ab} = e_n(\mathcal{H}_{ab})$ is unique and $(n-2)$-dimensional because e_n is a homeomorphism, and dimension is a topological invariant. Then, since aff $\Delta \subset \mathcal{R}^n$ is bounded away from 0, the latter is independent of every element of $\mathcal{H}'_{ab}$. Therefore, $\mathcal{I}_{ab} := \text{aff}(\mathcal{H}'_{ab} \cup \{0\})$ is an $(n-1)$-dimensional linear subspace of $\mathcal{R}^n$.

For the latter part, we note that if $\mathcal{H}_{ab} \cap \mathcal{B}_{ab}$ is nonempty, then because $\mathcal{I}_{ab} \cap \Delta = \mathcal{H}'_{ab} \cap \Delta$, $\mathcal{H}'_{ab}$ must also have nonempty intersection with $\mathcal{B}_{ab}$. Then, if $p' \in \mathcal{H}'_{ab} \cap \mathcal{B}_{ab}$, by corollary (1.1), it must be that $e_n^{-1}(p') \in \mathcal{B}_{ab}$ if and only if $e_n(p) \in \mathcal{B}_{ab}$. This means that $\mathcal{H}_{ab} \cap \mathcal{B}_{ab}$ is nonempty which of course contradicts proposition (2.1) together with the definition of $\mathcal{H}_{ab}$. Since the same is also true of $\mathcal{W}_{ab}$ by symmetry, the proof is complete. □

Remark 2.1. *For each $a, b \in A$, we define the subspace $\mathcal{I}^{\perp}_{ab}$ to be the orthogonal complement of $\mathcal{I}_{ab}$ in $\mathcal{R}^n$. Also, we denote the open half-space in $\mathcal{R}^n$ that is determined by $\mathcal{I}_{ab}$ and contains $\mathcal{B}_{ab}$ by $\mathcal{Y}_{ab}$. Then by the fundamental theorem of linear algebra, $\mathcal{I}^{\perp}_{ab}$ is of dimension 1.*

We now present the first representation result that consists of an order homeomorphism between $(A, \succ_p)_{p\in\Delta}$ and $(\mathcal{R}, >)$.

Theorem 2.1. *Preferences $(A, \succ_p) : p \in \Delta$ satisfy conditions (Asy.), (WP), (C'ty) and (p-SB) if and only if for each pair $a, b \in A$, there exists $f^{ab} : S \to \mathcal{R}$ and $F_{ab} : \Delta \to \mathcal{R}$ such that for all $p \in \Delta$ the following three properties hold*

i. $F_{ab}(p) = \sum_s p_s f^{ab}_s$,

ii. $F_{ba} = -F_{ab}$, *and*

iii. $a \succ_p b \Leftrightarrow F_{ab}(p) > 0$,

Moreover, if $\mathcal{B}_{ab}$ *and* $\mathcal{W}_{ab}$ *are both nonempty, then up to a positive scalar multiple,* f_{ab} *is unique.*

From here on we will refer to this kind of representation as an *expected preference representation*, and F_{ab} will be referred to as an *expected preference operator*.

Proof. As usual the necessity of the conditions is immediate with (Asy.) following from (ii) and (iii); (WP) and (p-SB) from (iii) and (i); and (C'ty) directly from (i). We now prove that the conditions are sufficient. Take any distinct pair $a, b \in A$. Let $f^{ab} : S \to \mathcal{R}$, be the function whose values are such that for each $s \in S$, f^{ab}_s is the s^{th} entry of a vector in $f_{ab} \in \mathcal{I}^{\perp}_{ab} \cap \mathcal{Y}_{ab}$. In the same way, we define the function f^{ba} in terms of the vector $-f_{ab}$.

Now the inner product in $\mathcal{R}^n$ of such a vector with any element of $\mathcal{B}_{ab} \subset \mathcal{Y}_{ab}$ is always positive, and $\langle f_{ab}, p \rangle$ is equal to $\sum_s p_s f^{ab}_s$. By letting $F'_{ab}(\cdot) := \langle f_{ab}, \cdot \rangle$, we define a bounded linear functional from $\mathcal{R}^n$ to $\mathcal{R}$, we can then let F_{ab} be defined as the restriction of F'_{ab} to Δ. On the other hand if p lies in $\Delta - \mathcal{B}_{ab}$ then either $p \in \mathcal{I}_{ab} \cap \Delta = \mathcal{N}_{ab}$ in which case $F_{ab}(p) = 0$ or $p \in \mathcal{W}_{ab}$. If $p \in \mathcal{W}_{ab}$, then $\mathcal{W}_{ab}$ is nonempty, and by proposition (2.1) we know that $\mathcal{I}_{ab}$ strictly separates $\mathcal{B}_{ab}$ from $\mathcal{W}_{ab}$. That is $\mathcal{W}_{ab}$ lies in the opposite half-space to $\mathcal{B}_{ab}$. Thus $F_{ab}(p) < 0$ as required. Now, note that because $f_{ba} = -f_{ab}$, the linear functional $F_{ba} := \langle f_{ba}, \cdot \rangle$ is equal to $-F_{ab}$, and it too satisfies properties i, ii and iii.

For any $a = b$, if we define $f_{aa} = 0 \in \mathcal{R}^n$, then F'_{aa} is the zero operator, with kernel equal to all of $\mathcal{R}^n$, and hence $F_{aa}(p) = 0$ for all $p \in \Delta$ and it is plain to see that F_{aa} satisfies properties i, ii and iii.

Now suppose that $G : (A \times A) \times \Delta$ is another expected preference representation of $\{(A, \succ_p) : p \in \Delta\}$. We wish to show that in the case where both $\mathcal{B}_{ab}$ and $\mathcal{W}_{ab}$ are nonempty, for some $\theta > 0$, $G = \theta F$.

By the assumption that it is an expected preference representation we know that for all $a, b \in A$ and $p \in \Delta$,

$$a \succ_p b \quad \Leftrightarrow \quad G_{ab}(p) > 0,$$

and for some $g^{ab} : S \to \mathcal{R}$,

$$G_{ab}(p) = \mathbb{E}_p(g^{ab}) = \sum_s g_s^{ab} \, p_s.$$

Therefore, for each a, b, the set of values $\{g_s^{ab}, s \in S\}$ uniquely defines a vector in $\mathcal{R}^n$, via the transformation

$$\{g_s^{ab}, s \in S\} \mapsto g_{ab} := \sum_s g_s^{ab} \, \delta_s,$$

where $\{\delta_s : s \in S\}$ is the standard basis of $\mathcal{R}^n$. This vector, in turn, uniquely defines a linear operator G'_{ab} on $\mathcal{R}^n$ such that both the element 0 and the set $\mathcal{N}_{ab}$ are contained in its kernel.

Now whenever $\mathcal{B}_{ab}$ and $\mathcal{W}_{ab}$ are both nonempty, by lemma (2.1), $\mathcal{N}_{ab}$ intersects the interior of Δ and it is $n-2$-dimensional. It therefore contains a set $\{p_1, \ldots p_{n-1}\}$ of $n-1$ affinely independent elements. The linear hull of this set equal to the linear subspace $\mathcal{I}_{ab}$ which is the kernel of F_{ab}. Now note that for any linear combination $\{\lambda_i\}$ of the $\{p_i\}$,

$$G'_{ab}(\lambda_1 p_1 + \cdots + \lambda_{n-1} p_{n-1}) = \sum_{i=1}^{n-1} \lambda_i G'_{ab}(p_i)$$

is equal to 0 because $G'_{ab}(p_i) = 0$ for all i. Therefore, $\ker G'_{ab} = \ker F'_{ab}$, so that $g_{ab} \in \mathcal{I}_{ab}^{\perp}$. Then the fact that for all $p \in \Delta$,

$$G_{ab}(p) > 0 \quad \Leftrightarrow \quad F_{ab}(p) > 0$$

implies that $g_{ab} \in \mathcal{Y}_{ab}$. This of course implies that $G_{ab} = \theta F_{ab}$ for some $\theta > 0$. $\square$

2.3 Conditions on triples of alternatives and separability across alternatives.

So far, as we have only imposed conditions on how preferences rank pairs of elements of A in isolation. Major differences with context-free preferences are revealed once we consider how the no-strict-preference sets $\{\mathcal{N}_{ab} : a, b \in A\}$ are arranged in relation to one another in context space. This we must do if we are to obtain a utility representation of preferences. Unlike the model of [vNM], with preferences defined on $\Delta(A \times S)$ where the level sets of the expected utility function are all parallel with one another, or the model of Dekel (1986) where level sets may fail to be parallel provided they do not intersect, in the context preferences approach, the sets $\{\mathcal{N}_{ab} : a, b \in A\}$ may well intersect one another without contradicting conditions such as transitivity a form of which we introduce below. Indeed, although it is not necessary for the existence of an expected utility function, if the intersection of these sets is not of maximal dimension then in general we will be unable say more about the representation than we have in the previous section.

The first property that is at the heart of the discussion in the previous paragraph is a subtle property of context preferences that have an expected utility representation that relates triples of elements of A. Following Bourbaki (1970, p.137) we will refer to it as the *groupoid property*. In the present setting, this property applies to the family of expected preference operators $\{F_{ab} : \Delta \to \mathcal{R}\}_{a,b \in A}$ that were introduced in the previous section, and is characterized as follows: for any a, b, b' and c in A if

$$F_{ab}(p) + F_{b'c}(p) = F_{ac}(p)$$

for each $p \in \Delta$, then $b \sim_p b'$ for each $p \in \Delta$. The groupoid property is closely related to "tradeoff consistency" as proposed by Kobberling and Wakker (2003) which takes the form $(a, b) \sim^*_{\cdot} (a, b')$ implies $b \sim_{\cdot} b'$, where $\sim^*$ is an order that is defined on differences (or more abstractly speaking simply pairs) and $\sim_{\cdot}$ is a standard context ordering. The following theorem shows that *context* preferences have an expected utility representation if and only if they have an expected preference representation satisfying the *strong groupoid property.* This is simply the groupoid property and its converse. The theorem will be instrumental in what follows.

Theorem 2.2. *For context preferences* $\{(A, \succ_p) : p \in \Delta\}$, *the following two statements are equivalent:*

(1) There exists an expected preference representation $F : (A \times A) \times \Delta \to \mathcal{R}$ *of preferences such that for each* a, b, b' *and* $c \in A$*, and all* $p \in \Delta$*,*

$$F_{ab}(p) + F_{b'c}(p) = F_{ac}(p)$$

if and only if $b \sim_p b'$ *for all* $p \in \Delta$*.*

(2) There exists an expected utility representation $U : A \times \Delta \to \mathcal{R}$ *of preferences.*

Moreover, the representation in (1) is unique up to a ratio scale (CRS) if and only if the representation in (2) is unique up to unit-comparability (CUC) across states.

Proof. First suppose statement (1) is true. Let $F_{ij}(p) = \sum_s p_s f_s^{ij}$ for $f^{ij} : S \to \mathcal{R}$, and let f_{ij} be the vector with elements $\{f_s^{ij}\}$. Then for all $i \in A$, and arbitrary a, define $u_i := f_{ia}$, so that for all $c, d \in A$, $p \in \Delta$ we have

$$\begin{aligned} F_{cd}(p) = &= \langle f_{cd}, p \rangle \\ &= \langle (f_{cb} + f_{b'd}), p \rangle \\ &= \langle f_{cb}, p \rangle + \langle f_{b'd}, p \rangle \\ &= \langle u_c, p \rangle - \langle u_d, p \rangle. \end{aligned}$$

Now if for each $i \in A$ and $s \in S$ we define u_s^i to be the s^{th} element of u_i, then $u_s^i \equiv f_s^{ia}$ and let $U : A \times \Delta \to \mathcal{R}$, $(a, p) \mapsto \sum_s p_s u_s^a$, then for all $c, d \in A$ and $p \in \Delta$,

$$c \succ_p d \quad \Leftrightarrow \quad F_{cd}(p) > 0 \quad \Leftrightarrow \quad U(c, p) > U(d, p).$$

Thus statement (1) implies statement (2). Note that we only needed the if part of the statement "if and only if $b \sim_p b'$ for all $p \in \Delta$".

Now suppose that statement (2) is true. That is there exists an expected utility representation $U : A \times \Delta \to \mathcal{R}$ of preferences $\{(A, \succ_p) : p \in \Delta\}$. By definition (1.2), there exists $u : A \times S \to \mathcal{R}$, with $U(a,p) := \sum_s p_s u_s^a$ and for all $c, d \in A$ and $p \in \Delta$

$$i \succ_p j \quad \Leftrightarrow \quad U(c,p) > U(d,p).$$

Now for each $i, j \in A$ and $p \in \Delta$, let $F_{ij}(p) := U(i,p) - U(j,p)$. Then clearly $(F_{ij}\,;\, i, j \in A)$ is an expected preference representation. Moreover, if for every $a, b, b', c \in A$ and $p \in \Delta$ we have

$$\begin{aligned} F_{ab}(p) + F_{b'c}(p) &= \sum_s p_s(u_s^a - u_s^b) + \sum_s p_s(u_s^{b'} - u_s^c) \\ &= F_{ac}(p), \end{aligned} \tag{2.1}$$

then since this holds for all degenerate probability measures δ_s, we have $u_s^a - u_s^b = u_s^a - u_s^{b'}$ and therefore $b \sim_p b'$ for all $p \in \Delta$ by condition (WP) which is itself necessary for (2). On the other hand, if $b \sim_p b'$ for all $p \in \Delta$, then equation (2.1) holds for all $p \in \Delta$ simply because U is a utility representation. Thus statement (2) implies statement (1).

We now show that if the expected utility representation is unique up to unit-comparability across states the corresponding expected preference representation is unique up to a ratio scale. We prove the contrapositive of this statement. To this end, suppose $(G_{ij}\,;\, i, j \in A)$ is another expected preference representation of $\{(A, \succ_p) : p \in \Delta\}$. Then by the first part of the existence proof above, there exists an expected utility representation $W : A \times \Delta \to \mathcal{R}$, $(a, s) \mapsto \sum_s p_s w_s^a$ such that for all $i, j \in A$ and $p \in \Delta$,

$$i \succ_p j \quad \Leftrightarrow \quad G_{ij}(p) > 0 \quad \Leftrightarrow W(i, p) > W(j, p).$$

If for some $c, d \in A$, the following functions from S to $\mathcal{R}$ satisfy (point wise)

$$w^c - w^d \equiv W(c, \cdot) - W(d, \cdot) \neq \theta\big(U(c, \cdot) - U(d, \cdot)\big) \equiv \theta(u^c - u^d)$$

for all $\theta > 0$. This is equivalent to the existence of an element $s \in S$ such that $w_s^c - w_s^d \neq \theta(u_s^c - u_s^d)$. From this we conclude that

$$w_s^c \neq \theta u_s^c + \kappa_s,$$

where $\kappa_s := -\theta u_s^d + w_s^d \in \mathcal{R}$. That is, the existence of an expected preference representation G that is not a positive scalar multiple of F implies that the expected utility representation is not unique up to unit-comparability.

We now show, also by proving the contrapositive statement, that if the expected preference representation F, that we defined in the first part of the existence proof above, is unique up to a ratio scale, then the expected utility representation that it gave rise to is unique up to unit-comparability. To this end, suppose that $V : A \times \Delta \to \mathcal{R}$, $(i, p) \mapsto \sum_s p_s v_s^i$ is another expected utility representation such that for some $i \in A$, the function $v^i : S \to \mathcal{R}$ pointwise satisfies $v^i \neq \theta u^i + \kappa$, for all $\theta > 0$ and all $\kappa S \to \mathcal{R}$. That is, the expected utility representation U that we have identified is not unique up to unit-comparability across states. Then, pointwise, this implies that

$$v^c - v^d \neq \theta u^c + \kappa - v^d$$

for all such θ and κ. Then letting $\kappa := -\theta u^d + v^d$, we see that for all $\theta > 0$,

$$v^c - v^d \neq \theta(u^c - u^d).$$

Now let G_{ij}, $i, j \in A$ be the respective expected preference representation that V gives rise to: that is $g^{ij} := v^i - v^j$ for every $i, j \in A$. Then for all $\theta > 0$ we have

$$G_{cd} \neq \theta F_{cd},$$

so that F is not unique upto a positive scalar multiplication, and hence not unique up to a ratio scale. □

Definition 2.4. *We will say that an expected preference representation has the* strong groupoid *property if it satisfies (1) of theorem (2.2).*

An immediate implication of the strong groupoid property is that preferences satisfy *negative-transitivity*. This is because preferences satisfy $F_{ij}(p) > 0$ if and only if $i \succ_p j$, and so by the strong groupoid property, we have that for all a, b and c in A, and all p in Δ, if $F_{ab}(p) \leqslant 0$ and $F_{bc}(p) \leqslant 0$ then $F_{ac}(p) \leqslant 0$ which is equivalent to

$$\neg(a \succ_p b) \;\wedge\; \neg(b \succ_p c) \quad \Rightarrow \quad \neg(a \succ_p c)$$

which is the definition of negative-transitivity which we denote by condition (NT).

Definition (Negative transitivity (NT)).

For all $a, b, c \in A$ and $p \in \Delta$: if $\neg(a \succ_p b)$ and $\neg(b \succ_p c)$, then $\neg(a \succ_p c)$; equivalently,

$$p \in \mathcal{B}_{ac} \quad \Rightarrow \quad p \in \mathcal{B}_{ab} \cup \mathcal{B}_{bc}\,.$$

Negative-transitivity, as the logical inverse of transitivity of strict preference, can be viewed as shorthand for a condition which states that all sequences involving $\succ_{\cdot}$ and $\sim_{\cdot}$ are transitive, with $\succ$ being the dominant relation in the sequence: for instance if $a \succ_{\cdot} b \sim_{\cdot} c$ then $a \succ_{\cdot} c$; and if $a \sim_{\cdot} b \sim_{\cdot} c$ then $a \sim_{\cdot} c$. It is well known that (NT) and (Asy) are together equivalent to assuming a weak preference relation $\gtrsim_{\cdot}$ is both complete and transitive. We choose this approach here, as we feel it is not only more appropriate, given the nature of (C'ty), but also because in future work we wish to weaken (NT) to transitivity so as to allow for of strict preference which, as we mentioned in chapter 1, seems to be a more intuitive and natural condition for *context* preferences.

As we discuss in example (2.2), condition (NT) together with the conditions on pairs introduced in the previous section are insufficient for the strong groupoid property. First we provide a counterexample to condition (NT).

Example 2.1. *[Getting to university continued] Recall from chapter 1, Val's context preferences $\{(A, \succ_p) : p \in \Delta\}$ with $A := \{a, b, c\}$, $S := \{r, s, t\}$, where $a \succ_s c \succ_s b$ when it is sunny, $c \succ_r b \succ_r a$ when it is raining and $b \succ_t a \succ_t c$ when it is icey. In this case, if conditions (Asy.), (C'ty), (WP), (p-SB) are satisfied, then theorem (2.1) ensures that there is an expected preference representation. If condition (NT) is not satisfied, the sets $\mathcal{N}_{ab}$, $\mathcal{N}_{ac}$ and $\mathcal{N}_{bc}$ might well be of the following form.*

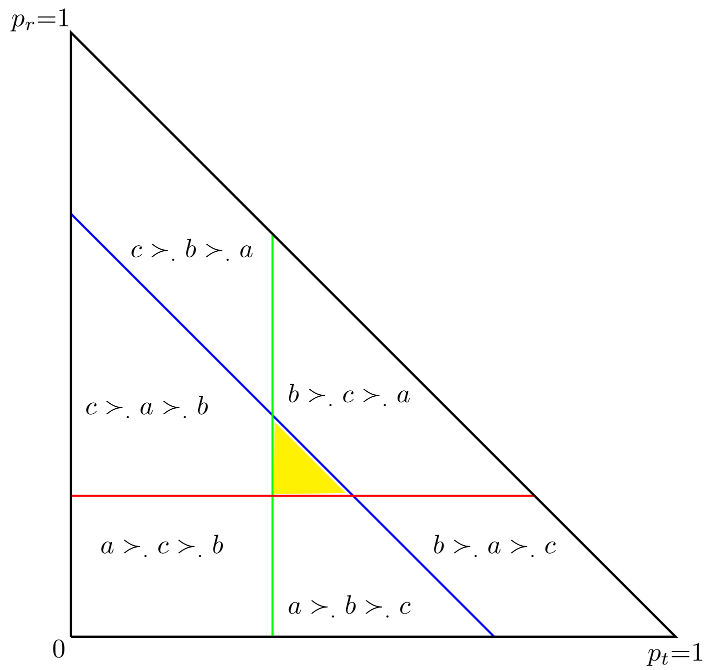


Figure 2.1: The shaded region in the center represents the region in which $a \succ_. b \succ_. c \succ_. a$. Such preferences may be explained by regret aversion as in Loomes and Sugden (1982).[3]

Suppose that Val's preferences do satisfy (NT), and moreover, using theorem (2.2), suppose an analyst is able to deduce that her preferences give rise to rock-paper-scissors type pay-off vectors $u_a := (0, -1, 1)$*,* $u_b := (1, 0, -1)$ *and* $u_c := (-1, 1, 0)$*. Then the matrix of payoff differences is:*

$$\begin{pmatrix} f_{ab}^T \\ f_{ac}^T \\ f_{bc}^T \end{pmatrix} = \begin{pmatrix} (u_a - u_b)^T \\ (u_a - u_c)^T \\ (u_b - u_c)^T \end{pmatrix} = \begin{pmatrix} -1 & -1 & 2 \\ 1 & -2 & 1 \\ 2 & -1 & -1 \end{pmatrix}$$

Thus all p *in* $\mathcal{N}_{ab}$ *satisfy* $0 = -p_r - p_s + 2p_t = 2 - 3p_r - 3p_t$*. Equivalently,* $\mathcal{N}_{ab} = \{p \in \operatorname{cl} \mathcal{R}^2_{>0} : p_r = \frac{2}{3} - p_t\}$*. Similarly,* $\mathcal{N}_{ac}$ *is characterized by* $\{p : p_r = \frac{1}{3}\}$

and finally, $\mathcal{N}_{bc}$ is equal to $\{p : p_t = \frac{1}{3}\}$. All three are plotted in the following figure.

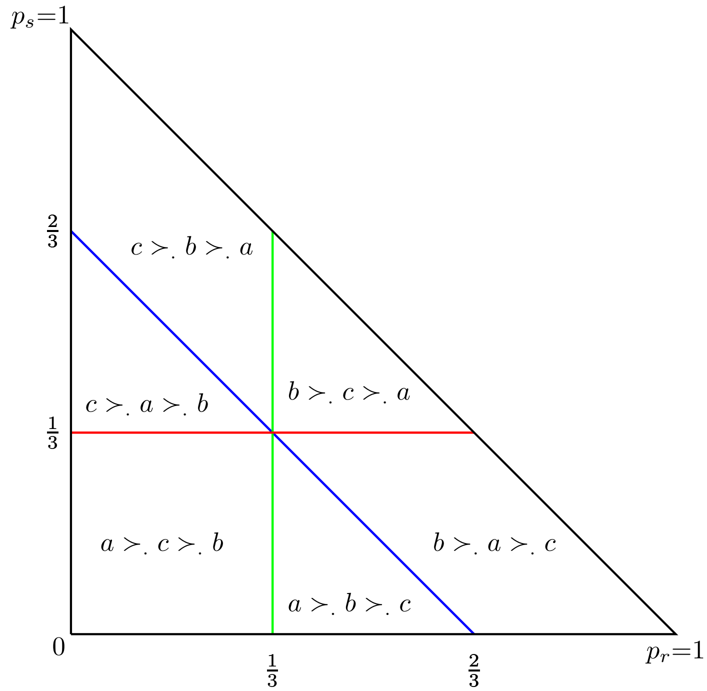


Figure 2.2: The case where Val's preferences are identical to those implied by the canonical rock-paper-scissors payoff matrix.

Example 2.2. *[Getting to university continued] Suppose instead that Val's preferences satisfy $c \succ_r a \succ_r b$, but are otherwise similar to example (2.1); in particular, condition (NT) and (p-SB) are satisfied. In this case, it may well be that preferences are like those in figure (2.2).*

As example (2.2) shows, without further conditions we are not able to improve upon the representation of theorem (2.1). We note that the example is not exceptional, indeed expected utility is the special case where for every triple $\{a, b, c\}$ of distinct elements in A, the hyperplanes determined by the sets $\mathcal{N}_{ij}$,

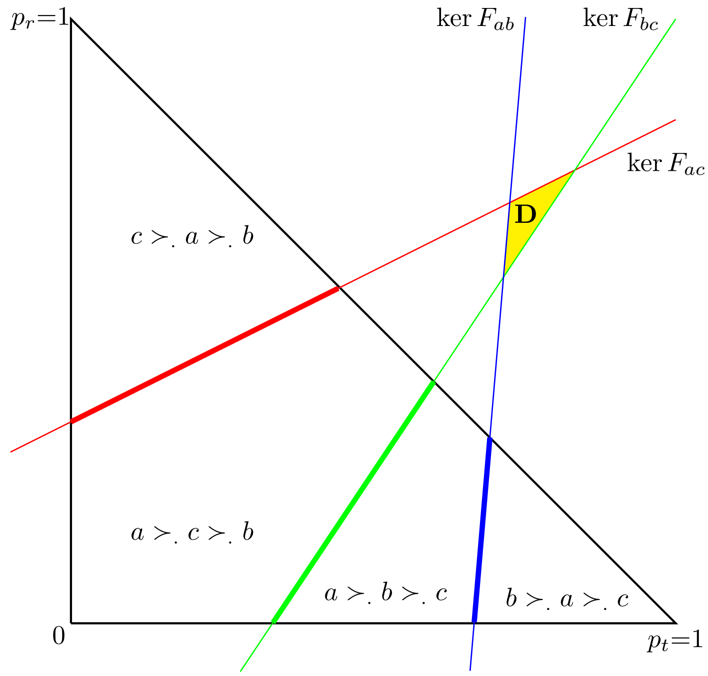


Figure 2.3: The case where preferences satisfy condition (NT), but where there is no expected utility representation. The reason is that for any x in the shaded region, labeled D, we have $F_{ab}(x) < 0$, $F_{bc}(x) < 0$ and $F_{ac}(x) > 0$ which clearly violates the strong groupoid property.

for $i \neq j$ in $\{a, b, c\}$ are congruent as opposed to lying in general position.

It is clear that more precise information regarding preferences is needed if in general we are to be assured of an EU representation. The question is how to do this in a minimal way, so as to impose the least restrictive conditions on the smallest possible domain for preferences. Extending preferences $(\succ_p, p \in \Delta)$ to the space of mixtures of elements of the space of alternatives A, which we denote by $\Delta(A)$ provides more information regarding the strength of preference for one alternative over another for each p, but it does not ensure consistency of preferences across p.

Definition (Div.)**.** *For any $d \in A$ let (a, b, c) be a list of distinct alternatives that excludes d. Preferences are* weakly diverse *if either:*

1) for any such list, there exists p such that $a \succ_p b \succ_p c \succ_p d$; or

2) for any such list, there exists p such that $d \succ_p a \succ_p b \succ_p c$

or both. If $|A| < 4$, then for any strict ordering of A, there is a p such that $\succ_p$ is that ordering.

The condition that appears in [GS03] as well as Gilboa and Schmeidler (1995, 2001) is similar to (Div.). It requires that for every possible list (a, b, c, d) of distinct alternatives there is a $p \in \Delta$ such that $a \succ_p b \succ_p c \succ_p d$. Clearly this implies that both (1) and (2) of definition (Div.) hold and as such the class of context preferences that satisfy (Div.) is substantially larger. We provide some support for this claim in the final section of this chapter. There we will also show that at least for the present context space, Δ, (Div.) is equivalent to the next condition.

Definition (Weak diversity (Div.))**.**

i) For any list (a, b, c) of distinct alternatives, there is a $p \in \Delta$ such that $a \succ_p b \succ_p c$. If $|A| = 2$, then for each strict ordering of A, there is a p such that $\succ_p$ is that ordering.

ii) For any list (a, b, c, d) of distinct alternatives, there exists p in Δ such that $a \sim_p b \sim_p c \not\succ_p d$.

Remark 2.2. *Note that* (Div.(ii)) *is void whenever* $|A| < 4$.

We now present the consequences of imposing (Div.(i)) along with (NT) and the pairwise conditions of the previous section.

Lemma 2.2. *Let* $|A| > 2$. *If context preferences* $(A, \succ_p)_{p \in \Delta}$ *satisfy conditions (Asy.), (WP), (C'ty), (p-SB), (NT) and (Div.(i)), then for all distinct* $a, b, c \in A$, $\mathcal{N}_{abc}$ *is nonempty and of dimension greater than or equal to* $n-3$.

Proof. Fix three distinct elements $a, b, c \in A$. First note that condition (NT) implies that $\mathcal{N}_{abc} = \mathcal{N}_{ab} \cap \mathcal{N}_{bc}$. By condition (Div.(i)), we know that

$$\mathcal{B}_{ab} \cap \mathcal{B}_{bc} := \{p \in \Delta : a \succ_p b \succ_p c\} \neq \varnothing,$$

and since $\mathcal{B}_{ca} \cap \mathcal{B}_{ab} \equiv \{c \succ_{\cdot} a \succ_{\cdot} b\}$ is nonempty and via (NT) contained in $\mathcal{B}_{ab} \cap \mathcal{W}_{bc}$, we see that the latter is also nonempty. Moreover, condition (Asy.) implies that $\mathcal{B}_{bc} \cap \mathcal{W}_{bc}$ is empty, so that $\mathcal{B}_{ab} \cap \mathcal{B}_{bc}$ and $\mathcal{B}_{ab} \cap \mathcal{W}_{bc}$ are disjoint subsets of $\mathcal{B}_{ab}$. By condition (C'ty), and the fact that the intersection of open sets is open, we know that these two subsets are also open in both Δ and $\mathcal{B}_{ab}$. By condition (p-SB), $\mathcal{B}_{ab}$ is convex and therefore connected. Then proposition (3.5) implies that $\mathcal{B}_{ab} \cap \mathcal{B}_{bc}$ and $\mathcal{B}_{ab} \cap \mathcal{W}_{bc}$ are (topologically) separated sets by their nonempty complement $\mathcal{B}_{ab} \cap \mathcal{N}_{bc}$ in $\mathcal{B}_{ab}$.[4]

By an identical argument the sets $\mathcal{W}_{ab} \cap \mathcal{B}_{bc}$ and $\mathcal{W}_{ab} \cap \mathcal{W}_{bc}$ are also open, nonempty, disjoint subset of $\mathcal{W}_{ab}$ and therefore topologically separated by the nonempty set $\mathcal{W}_{ab} \cap \mathcal{N}_{bc}$ in $\mathcal{W}_{ab}$. Now $\mathcal{B}_{ab}$ and $\mathcal{W}_{ab}$ are themselves (topologically) separated sets, so that $\mathcal{B}_{ab} \cap \mathcal{N}_{bc}$ and $\mathcal{W}_{ab} \cap \mathcal{N}_{bc}$ are, by definition (3.7),

[4]Please see the next chapter for the definition of topologically separated sets

(topologically) separated sets.

We will show that $\mathrm{ri}(\mathcal{N}_{bc})$ also has nonempty intersection with both $\mathcal{B}_{ab}$ and $\mathcal{W}_{ab}$. Suppose that $\mathrm{ri}(\mathcal{N}_{bc}) \subset \mathcal{B}_{ab}$. Then since since $\mathcal{N}_{bc}$ is by proposition (2.1), a convex subset of a Euclidean space, theorem (2.3.8) of Webster p. 64 ensures that $\mathrm{cl}(\mathcal{N}_{bc}) = \mathrm{cl}(\mathrm{ri}\,\mathcal{N}_{bc})$, which is equal to $\mathcal{N}_{bc}$ as the latter is closed in Δ, and Δ is closed. However, this implies that $\mathrm{cl}\,\mathcal{B}_{ab}$ contains $\mathcal{N}_{bc}$, so that by the definition of (topologically) separated sets we must have $\mathcal{W}_{ab} \cap \mathcal{N}_{bc} = \varnothing$: a contradiction of what we have shown in the previous paragraph.

Furthermore, the set $\mathrm{ri}\,\mathcal{N}_{bc}$, as the interior of a convex set is, by theorem (2.3.5) of Webster p.62, also convex and therefore connected. Then since $\mathcal{B}_{ab} \cap \mathrm{ri}\,\mathcal{N}_{bc}$ and $\mathcal{W}_{ab} \cap \mathrm{ri}\,\mathcal{N}_{bc}$ are open, nonempty, and disjoint in $\mathrm{ri}\,\mathcal{N}_{bc}$, it must be that their complement in $\mathrm{ri}\,\mathcal{N}_{bc}$, $\mathcal{N}_{ab} \cap \mathrm{ri}\,\mathcal{N}_{bc}$, separates them. This shows that $\mathcal{N}_{ab} \cap \mathrm{ri}\,\mathcal{N}_{bc}$ is nonempty and it disconnects the open subset $\mathrm{ri}\,\mathcal{N}_{bc}$ of $\mathcal{H}_{bc}$.

Now since $\mathcal{H}_{bc}$ is a hyperplane in $\mathcal{R}^{n-1}$ it is homeomorphic to $\mathcal{R}^{n-2}$, so that $\mathrm{ri}\,\mathcal{N}_{bc}$ is an $n-2$-dimensional manifold. Then corollary (1) of theorem (IV 4) of Hurewicz and Wallman p.48 states that a manifold such as $\mathrm{ri}\,\mathcal{N}_{bc}$ cannot be disconnected by a subset of dimension less than or equal to $n-4$. This implies that the dimension of $\mathcal{N}_{ab} \cap \mathrm{ri}\,\mathcal{N}_{bc}$, and hence $\mathcal{N}_{abc}$, is greater than or equal to $n-3$. $\square$

Lemma 2.3. *Let $|A| > 2$. If conditions (Asy.), (WP), (C'ty), (p-SB), (NT) and (Div.(i)) hold for context preferences $\{(A, \succ_p) : p \in \Delta\}$, then for all distinct*

a, b and c in A,

$$\dim(\mathcal{N}_{abc}) = \dim \mathcal{H}_{abc} = \dim(\mathcal{H}_{ab} \cap \mathcal{H}_{bc}) = n - 3,$$

where $\mathcal{H}_{abc}$ denotes the intersection $\mathcal{H}_{ab} \cap \mathcal{H}_{bc} \cap \mathcal{H}_{ac}$.

Proof. Let a, b and c be distinct elements in A. By proposition (2.1) the affine hulls $\mathcal{H}_{ab}$ of $\mathcal{N}_{ab}$, $\mathcal{H}_{bc}$ of $\mathcal{N}_{bc}$ and $\mathcal{H}_{ac}$ of $\mathcal{N}_{ac}$ are all $n-2$-dimensional. By lemma (2.2), we know that the affine hull of $\mathcal{N}_{abc}$ is at least of dimension $n-3$. Then because the affine hull of a set in $\mathcal{R}^{n-1}$ is defined to be the intersection of all the flats in $\mathcal{R}^{n-1}$ that contain the set, and because we know that for all $i \neq j$ in $\{a, b, c\}$, $\mathcal{N}_{abc} \subset \operatorname{aff} \mathcal{N}_{ij} = \mathcal{H}_{ij}$ simply because $\mathcal{N}_{abc} \subset \mathcal{N}_{ij}$, we conclude that $\operatorname{aff} \mathcal{N}_{abc} \subset \mathcal{H}_{abc}$ and that therefore the dimension of $\mathcal{H}_{abc}$ is greater than or equal to $n-3$. In particular, this implies that $\mathcal{H}_{ab} \cap \mathcal{H}_{bc}$ is nonempty, so that by theorem 1.3.8 of Webster p.14, we find that

$$\begin{aligned}\dim(\mathcal{H}_{ab} \cap \mathcal{H}_{bc}) &= \dim \mathcal{H}_{ab} + \dim \mathcal{H}_{bc} - \dim(\mathcal{H}_{ab} + \mathcal{H}_{bc}) \\ &= n - 2 + n - 2 - (n - 1) = n - 3,\end{aligned}$$

where $\dim(\mathcal{H}_{ab} + \mathcal{H}_{bc}) = n - 1$ because $\mathcal{H}_{ab} \neq \mathcal{H}_{bc}$ by condition (Div.(i)). That is, if $\mathcal{H}_{ab} = \mathcal{H}_{bc}$ then $\mathcal{B}_{ab}$ is strictly separated from either $\mathcal{B}_{bc}$ or $\mathcal{W}_{bc}$, so that either $\{a \succ_{.} b \succ_{.} c\} = \mathcal{B}_{ab} \cap \mathcal{B}_{bc}$ or $\{c \succ_{.} a \succ_{.} b\} \subset \mathcal{B}_{ab} \cap \mathcal{W}_{bc}$ is empty.

In summary, the above, together with lemma (2.2), allows us to conclude that

$$n - 3 \leqslant \dim(\operatorname{aff} \mathcal{N}_{abc}) \leqslant \dim \mathcal{H}_{abc} \leqslant \dim(\mathcal{H}_{ab} \cap \mathcal{H}_{bc}) = n - 3.$$

$\square$

Next we briefly state and prove a result about the relationship between half-spaces and cones generated by intersections of half-spaces with the sphere, that will be used often in what follows.

Proposition 2.2. *Let $\{X_\alpha\}$ be a family of open half-spaces in a Euclidean space, E, such that 0 lies in the boundary of each member and let $X = \bigcap_\alpha X_\alpha$. Then $X = \operatorname{cone}(\mathcal{S} \cap X_\alpha) - \{0\}$, where $\mathcal{S}$ is the boundary of an open ball of fixed radius, and is centered around 0.*

Proof. Note that if Y is an open half-space, with 0 in its boundary if and only if there exists an inward pointing vector f of Y. That is f such that $\langle f, x\rangle > 0$ for all $x \in Y$. Now suppose that x lies in

$$\operatorname{cone}(\mathcal{S} \cap Y) - \{0\} = \{y \in E : y = \lambda z,\, z \in \mathcal{S} \cap Y,\, \lambda > 0\}.$$

Then then there exists $z \in \mathcal{S} \cap Y$, $\lambda > 0$ such that $x = \lambda z$. Then since $z \in Y$, we know that $\langle f, z\rangle > 0$, then by linearity of inner product in each of its arguments, $\langle f, \lambda z\rangle > 0$, so that $x \in Y$. On the other hand, if $x \in Y$, then letting $\lambda = 1/\|x\|$, which is well defined and positive since 0 lies outside Y, we see that $\lambda x \in \mathcal{S} \cap Y$ and so x lies in $\operatorname{cone}(\mathcal{S} \cap Y) - \{0\}$. Indeed this is also true of arbitrary intersections of open half-spaces: the above argument holds with X replacing Y, with the statements holding simultaneously for every $\{f_\alpha\}$ in the corresponding family of inward pointing vectors. $\square$

The following consequence of (Div.(i)) is mentioned but not proved in the conclusion of Ashkenazi and Lehrer (2001). On the one hand it provides useful information about the set $\mathcal{N}_{abc}$, but it will also allow for a simple (almost diagrammatic) proof of the existence of a utility function for triples of elements of A.

Proposition 2.3. *Let $|A| > 2$. If context preferences $\{(A, >_p) : p \in \Delta\}$ satisfy (Asy.), (WP), (C'ty), (p-SB), (NT) and (Div.(i)), then for all distinct $a, b, c \in A$, $\mathcal{N}_{abc} \cap \operatorname{ri}\Delta$ is nonempty.*

Proof. Let a, b and c be distinct elements in A. We will first prove the following claim.

Claim. *If $\mathcal{N}_{abc} \subset \operatorname{rebd}\Delta$, then there exists $s \in S$ such that for all $p \in \mathcal{N}_{abc}$, $s \notin \operatorname{supp}(p)$.*

Proof of Claim. Recall that $\operatorname{ri}\Delta \equiv \{p \in \Delta : p_s > 0, \forall s \in S\}$, that is the set of probability measures such that $\operatorname{supp}(p) = S$. By the premise of the claim we know that for no $p \in \mathcal{N}_{abc}$ do we have $\operatorname{supp}(p) = S$, and if the conclusion of the claim is false, then

$$\bigcup_{p \in \mathcal{N}_{abc}} \operatorname{supp}(p) = S.$$

So let p_1 be an arbitrary element of $\mathcal{N}_{abc}$, and define S_1 to be its support. Then $1 \leqslant |S_1| < n$, so there exists $p_2 \in \mathcal{N}_{abc}$ such that $\operatorname{supp}(p_2)\backslash S_1 \neq \varnothing$. Now recall that as the intersection of convex sets, $\mathcal{N}_{abc}$ is convex, this implies that $r_2 := p_2 \frac{1}{2} p_1 \in \mathcal{N}_{abc}$, and moreover, $S_2 := \operatorname{supp}(r_2) = S_1 \cup \operatorname{supp}(p_2)$. The hypothesis that $\mathcal{N}_{abc} \subset \operatorname{rebd}\Delta$ then implies that $S_2 \neq S$, so that because S_2 must have at least one more element than S_1, we have $2 \leqslant |S_2| < n$.

We now make the induction hypothesis, and take r_j to be the j^{th} element in a sequence in $\mathcal{N}_{abc}$ that is recursively constructed as we did for r_2. That is

$$r_j = p_j\, 1/2\, r_{j-1} = p_j\, 1/2\, (p_{j-1}\, 1/2\, r_{j-2}) = \dots,$$

where for $i = 1, \dots, j$, $p_i \in \mathcal{N}_{abc}$, and $S_i := \operatorname{supp}(r_i)$ and for $i > 2$, $\operatorname{supp}(p_i)\backslash S_{i-1} \neq \varnothing$. Then as $r_j \in \operatorname{rebd}\Delta$, $j \leqslant |S_j| < n$, so that as in the first step, there exists $p_{j+1} \in \mathcal{N}_{abc}$ such that $\operatorname{supp}(p_{i+1})\backslash S_i \neq \varnothing$. Then once again convexity of $\mathcal{N}_{abc}$ implies that it contains $r_{j+1} := p_{j+1}\, 1/2\, r_j$, so that $S_{j+1} := \operatorname{supp}(r_{j+1}) \neq S$, and $j + 1 \leqslant |S_{j+1}| < n$.

Now as this completes the inductive step, there exists such an r_j for each $j \in \mathbb{N}$, and so it is also true for $j = n$. This provides the desired contradiction of the hypothesis that the both $\mathcal{N}_{abc} \subset \operatorname{rebd}\Delta$ and $\bigcup_{p \in \mathcal{N}_{abc}} \operatorname{supp}(p) = S$ can simultaneously hold. $\square$

Now recall that every $n - 2$-dimensional face of Δ is the convex hull of $n - 1$ extremal elements of Δ. By the above claim, we know that if $\mathcal{N}_{abc} \subset \operatorname{rebd}\Delta$, then $\mathcal{N}_{abc}$ is a subset of at least one such face Δ_{n-1} of Δ. Let $J := \operatorname{aff}\Delta_{n-1}$, and let Y and Z be the open half spaces in $\mathcal{R}^{n-1}$ determined by J. Since $J \cap \operatorname{ri}\Delta$ is empty, J supports Δ, and as such we may suppose that $\operatorname{ri}\Delta \subset Y$. Then as $\mathcal{N}_{abc} \subset \Delta_{n-1}$, $\operatorname{aff}\mathcal{N}_{abc}$ is a subset of J, so that by lemma (2.3)

$$\dim(\mathcal{H}_{ab} \cap \mathcal{H}_{bc} \cap J) = \dim(\mathcal{H}_{abc}) = n - 3.$$

Also, proposition (2.1) together with condition (Div.) implies that both $\mathcal{H}_{ab}$ and $\mathcal{H}_{bc}$ are distinct from J as they both intersect the relative interior of Δ.

Now consider the quotient of the affine hull of Δ, $\mathcal{R}^{n-1}$, with respect to $\mathcal{H}_{abc}$:

$$\mathcal{R}^{n-1}/\mathcal{H}_{abc} := \{y + \mathcal{H}_{abc} : \ y \in \mathcal{R}^{n-1}\}.$$

For all subsets X of $\mathcal{R}^{n-1}$, we denote the quotient with respect to $\mathcal{H}_{abc}$ by $\widetilde{X}$, and we will also denote an element of $\widetilde{\mathcal{R}^{n-1}}$ by $\widetilde{y}$, bearing in mind the fact that this is an $n-3$-dimensional subset of $\mathcal{R}^{n-1}$. By theorem 1 of Halmos (1974) p34, if X and $\mathcal{H}_{abc}$ are complementary subspaces in $\mathcal{R}^{n-1}$, then the correspondence that assigns to each vector y in X, the coset $y + \mathcal{H}_{abc}$ is an isomorphism between X and $\widetilde{\mathcal{R}^{n-1}}$. A natural example of such a subspace is the orthogonal complement of $\mathcal{H}_{abc}$ in $\mathcal{R}^{n-1}$:

$$\mathcal{H}_{abc}^{\perp} := \{x \in \mathcal{R}^{n-1} : \ \langle x, z \rangle = 0 \text{ for all } z \in \mathcal{H}_{abc}\},$$

By theorem 2 in the same section of Halmos, we know that $\widetilde{\mathcal{R}^{n-1}}$ is a 2-dimensional space.

Now for any hyperplane H in $\mathcal{R}^{n-1}$ that contains $\mathcal{H}_{abc}$, there exists $p \in H - \mathcal{H}_{abc}$, so that because $\widetilde{H \cap \mathcal{H}_{abc}} = \widetilde{\mathcal{H}_{abc}} = \widetilde{0}$, and $p \notin \widetilde{0}$, we conclude that $\text{aff}(\widetilde{0}, \widetilde{p})$ is a 1-dimensional subspace of $\widetilde{\mathcal{R}^{n-1}}$ that lies in $\widetilde{H}$. Now the fact that $\widetilde{H}$ is in fact equal to $\text{aff}(\widetilde{0}, \widetilde{p})$ follows from the fact that if there existed $\widetilde{y} \in \widetilde{H}$ such that $\widetilde{y} \notin \text{aff}(\widetilde{0}, \widetilde{p})$, the affine hull of these three independent points would be equal to the 2-dimensional space $\widetilde{\mathcal{R}^{n-1}}$. Then this would imply that $\widetilde{\mathcal{R}^{n-1}}$ is equal to $\widetilde{H}$, that is $\mathcal{R}^{n-1} = H + \mathcal{H}_{abc} = H$, which would contradict the assumption that H is a hyperplane in $\mathcal{R}^{n-1}$.

The above argument implies that $\widetilde{\mathcal{H}_{ab}}$, $\widetilde{\mathcal{H}_{bc}}$ and $\widetilde{J}$ are all one-dimensional subspaces of $\widetilde{\mathcal{R}^{n-1}}$. Indeed since $\mathcal{H}_{ab}$, $\mathcal{H}_{bc}$ and J are distinct hyperplanes with the set $\mathcal{H}_{abc}$ in common, they each contain an element that lies outside both of the others, and therefore outside $\mathcal{H}_{abc}$. We therefore conclude that $\widetilde{\mathcal{H}_{ab}}$, $\widetilde{\mathcal{H}_{bc}}$ and $\widetilde{J}$ are distinct one-dimensional subspaces of $\widetilde{\mathcal{R}^{n-1}}$.

Now let $\mathcal{S}$ be the unit sphere in $\widetilde{\mathcal{R}^{n-1}}$ that τ-separates the relative interior of the closed unit disc $\mathcal{D}$ centered around $\widetilde{0}$ from $\widetilde{\mathcal{R}^{n-1}} - \mathcal{D}$. Consider the intersection of $\mathcal{S}$ with any one-dimensional subspace H of $\widetilde{\mathcal{R}^{n-1}}$. As is the case in $\mathcal{R}^2$, because H intersects $\widetilde{0}$, it intersects $\mathcal{D}$, then since H is unbounded it cannot be contained in $\mathcal{D}$. Then since H is a subspace, it is connected, and so the open, nonempty and disjoint sets $H \cap \operatorname{ri} \mathcal{D}$ and $H \cap (\widetilde{\mathcal{R}^{n-1}} - \mathcal{D})$ are, by proposition (3.5), separated by $\mathcal{S}$ in H. This implies that $\mathcal{S} \cap H$ contains at least one point x. Now since $-x$ also lies on the boundary of the unit disc, and because H is a subspace, $-x$ also lies in $\mathcal{S} \cap H$. Indeed, since H is one-dimensional, $\operatorname{aff}(-x, x) = H$. Moreover, for every $0 \leqslant \lambda < 1$, we have $\lambda x \in \operatorname{ri} \mathcal{D}$, whilst for $\lambda > 1$, we have $\lambda x \notin \mathcal{D}$. Then since the same is true of $-x$, we conclude that $\mathcal{S} \cap H$ is equal to the pair of what are called antipodal points $\{-x, x\}$.

Now note that if H' is another such one-dimensional subspace of $\widetilde{\mathcal{R}^{n-1}}$, distinct from H, then because $H \cap H' = \widetilde{0}$, there exists $y \neq x$ such that $H' \cap \mathcal{S} = \{-y, y\}$. This implies that the intersection of $\mathcal{S}$ with each of the sets $\widetilde{\mathcal{H}_{ab}}$, $\widetilde{\mathcal{H}_{bc}}$ and $\widetilde{J}$ is characterized by a distinct pair of antipodal points.

Next we define $R : \widetilde{\mathcal{R}^{n-1}} \times [0, 2\pi] \longrightarrow \widetilde{\mathcal{R}^{n-1}}$ to be the standard (anti-clockwise) rotation matrix in $\mathcal{R}^2$.[5]

$$(x,\theta) \longmapsto R(x,\theta) = \begin{bmatrix} x_1\cos(\theta) - x_2\sin(\theta) \\ x_1\sin(\theta) + x_2\cos(\theta) \end{bmatrix}$$

Take x to be the element of $\widetilde{J}\cap\mathcal{S}$ satisfying the property that for all $0 < \theta < \pi$, $R(x,\theta) \in \widetilde{Y} \cap \mathcal{S}$, where we recall that Y is the open half-space in $\mathcal{R}^{n-1}$ that is determined by J and contains $\operatorname{ri}\Delta$. This we can do because $R(x,0)$ and $R(x,\pi)$ lie in $\widetilde{J}$, and by proposition (2.2) $\widetilde{Y} = \operatorname{cone}(\widetilde{Y} \cap \mathcal{S}) - \widetilde{0}$.

Now let y be an element of $\widetilde{\mathcal{H}_{ab}} \cap \mathcal{S}$. Then because $\widetilde{\mathcal{H}_{ab}}$ is distinct from $\widetilde{J}$, it holds that for some unique $0 < \theta < 2\pi$ other than π, we have $R(x,\theta) = y$. If $\theta < \pi$, then $y \in \widetilde{Y}$, and if not, then $\theta > \pi$ and since modulo 2π, $\theta + \pi$ is less than π, we see that $-y = R(x, \theta + \pi)$ is an element of $\widetilde{Y}$ instead. So without loss of generality there exists $0 < \theta_{ab} < \pi$, such that $y_{ab} = R(x, \theta_{ab})$ lies in $\mathcal{H}_{ab} \cap (\widetilde{Y} \cap \mathcal{S})$. Similarly, there exists $0 < \theta_{bc} < \pi$, such that $y_{bc} = R(x, \theta_{bc})$ is in $\mathcal{H}_{bc} \cap (\widetilde{Y} \cap \mathcal{S})$, and since $\widetilde{\mathcal{H}_{bc}}$ is distinct from $\widetilde{\mathcal{H}_{ab}}$ we know that $\theta_{bc} \neq \theta_{ab}$.

Without loss of generality, suppose that $\theta_{ab} < \theta_{bc}$, and denote the open set $]\theta_{ab}, \theta_{bc}[$ by Θ, and the open set $]\theta_{ab} + \pi, \theta_{bc} + \pi[$ by $\pi + \Theta$. Then note that $R(x,\Theta) \subset \widetilde{Y}$ and $R(x, \pi + \Theta) \subset \widetilde{Z}$, that is $\widetilde{J}$ strictly separates the two sets. Similarly, since Θ is a subset of $]\theta_{ab}, \theta_{ab} + \pi[$, and $\pi + \Theta$ is a subset of $]\theta_{ab} + \pi, 2\pi] \cup [0, \theta_{ab}[$, we see that $\widetilde{\mathcal{H}_{ab}}$ also strictly separates $R(x,\Theta)$ from $R(x, \pi + \Theta)$. An identical argument also shows that $\widetilde{\mathcal{H}_{bc}}$ also strictly separates the two sets.

[5]This a reduced form of the actual matrix we need, which is formally constructed in the main result that follows.

Recall that $\mathcal{Y}_{..}$ is the half-space in $\mathcal{R}^{n-1}$ that is determined by $\mathcal{H}_{..}$ and contains $\mathcal{B}_{..}$ and $\mathcal{Z}_{..}$ is the open half-space of $\mathcal{R}^{n-1}$ that is determined by $\mathcal{H}_{..}$ and contains $\mathcal{W}_{..}$. The definition of strict separation implies that if $R(x, \Theta)$ is a subset of say $\widetilde{\mathcal{Y}_{ab}} \cap \widetilde{\mathcal{Y}_{bc}}$, then $R(x, \pi + \Theta) \subset \widetilde{\mathcal{Z}_{ab}} \cap \widetilde{\mathcal{Z}_{bc}}$, or alternatively, if $R(x, \Theta)$ is a subset of $\widetilde{\mathcal{Y}_{ab}} \cap \widetilde{\mathcal{Z}_{bc}}$ then $R(x, \pi + \Theta)$ is a subset of $\widetilde{\mathcal{Z}_{ab}} \cap \widetilde{\mathcal{Y}_{bc}}$. In fact because the other two cases are obtained by replacing $\mathcal{Y}$ with $\mathcal{Z}$ in the preceding sentence, we will consider only this pair.

Consider the case where $R(x, \pi + \Theta) \subset \widetilde{\mathcal{Z}_{ab}} \cap \widetilde{\mathcal{Z}_{bc}}$. The fact that $R(x, \pi + \Theta) = \widetilde{\mathcal{Z}_{ab}} \cap \widetilde{\mathcal{Z}_{bc}} \cap \mathcal{S}$ follows from the fact that $R(x, \zeta)$ is an element of $\widetilde{\mathcal{Z}_{ab}} \cap \widetilde{\mathcal{Z}_{bc}} \cap \mathcal{S}$ if and only if $\theta_{ab} + \pi < \zeta < \theta_{bc} + \pi$. Then proposition (2.2) implies that $\widetilde{\mathcal{Z}_{ab}} \cap \widetilde{\mathcal{Z}_{bc}} \subset \widetilde{Z}$, so that because $\widetilde{Z} := \{y + \mathcal{H}_{abc} : y \in Z\}$, we immediately see that $\mathcal{Z}_{ab} \cap \mathcal{Z}_{bc}$ is a subset of Z. Then because $\mathcal{W}_{ab} \cap \mathcal{W}_{bc}$ is nonempty by condition (Div.), and open in Δ by condition (C'ty), it has nonempty relative interior, and as such, it contains elements of the relative interior of Δ. Moreover, because $\mathcal{W}_{ab} \cap \mathcal{W}_{bc}$ is a subset of $\mathcal{Z}_{ab} \cap \mathcal{Z}_{bc}$ there exists $y \in \operatorname{ri} \Delta$ such that $y \in Z$: a contradiction of the fact that $\operatorname{ri} \Delta \subset Y$.

The case where $R(x, \pi + \Theta) \subset \widetilde{\mathcal{Z}_{ab}} \cap \widetilde{\mathcal{Y}_{bc}}$ differs only in that we additionally need to show that the set $\mathcal{W}_{ab} \cap \mathcal{W}_{bc}$ is nonempty. This follows from the fact that it is equal to $\{b \succ_{.} a\} \cap \{b \succ_{.} c\}$, so that $\{b \succ_{.} a \succ_{.} c\}$ is its subset, together with condition (Div.). $\square$

We now begin the actual construction of the representation in the following lemma.

For all $a, b, c \in A$, we define $\mathcal{H}''_{abc}$, $\mathcal{H}'_{abc}$ and $\mathcal{H}_{abc} := \text{aff}(\mathcal{H}'_{abc} \cup \{0\})$, and by lemma (2.3) we know that $\dim(\mathcal{H}''_{abc}) = n - 3$, so that by an identical argument $\dim(\mathcal{H}_{abc}) = n - 2$. Furthermore, if $\mathcal{N}''_{abc} \cap \text{ri}\,\Delta''$, then we know that $\mathcal{N}''_{abc}$ is $n-3$-dimensional, so that $\mathcal{H}_{ab} \cap \Delta'$ is of dimension $n - 2$.

Lemma 2.4. *Let $A = \{a, b, c\}$. Let preferences $(A, \succ_p)_{p \in \Delta}$ satisfy condition (Div.(i)). Then the following two statements are equivalent:*

1. *$(A, \succ_p)_{p \in \Delta}$ also satisfy (Asy.), (WP), (C'ty), (p-SB) and (NT),*
2. *There exists a triple of functions F_{ab}, F_{bc} and F_{ac}, from Δ to $\mathcal{R}$, that that satisfy theorem (2.1), and have the property that for all p in Δ,*

$$F_{ab}(p) \; + \; F_{bc}(p) \; = \; F_{ac}(p).$$

Note that if statement 2 is satisfied, then by theorem (2.1) conditions (Asy.), (WP), (C'ty) and (p-SB) are all satisfied whether or not condition (Div.(i)) holds. Moreover, for all $p \in \Delta$, if $F_{ab}(p) \leqslant 0$ and $F_{bc}(p) \leqslant 0$, then by statement 2, $F_{ac}(p)$ is also less than or equal to zero. By theorem (2.1), this fact is equivalent to the statement $\neg(a \succ_p b)$ and $\neg(b \succ_p c)$ implies $\neg(a \succ_p c)$, which is the definition of negative transitivity (NT). This shows that 2 implies 1. The substance of the proof that follows is how in the presence of condition (Div.(i)), 1 is sufficient for 2.

Proof. Step 1. Throughout this proof we work in $\mathcal{R}^n$. By lemma (1.1) all the results we have obtained in this section in $\mathcal{R}^{n-1} = \text{aff}\,\Delta_0$ also hold in $\text{aff}\,\Delta \subset \mathcal{R}^n$. By lemma (2.1), $\mathcal{I}_{..}$ is equal to $\text{aff}(\mathcal{N}_{..}, 0)$ which is equal to a hyperplane through the origin in $\mathcal{R}^n$ that separates $\mathcal{B}_{..} \subset \mathcal{Y}_{..}$ from $\mathcal{W}_{..} \subset \mathcal{Z}_{..}$, where $\mathcal{Y}_{..}$ and $\mathcal{Z}_{..}$ are the half-spaces in $\mathcal{R}^n$ that are determined by $\mathcal{I}_{..}$. An identical argument shows that $\text{aff}(\mathcal{N}_{abc}, 0) = \text{aff}(\mathcal{H}_{abc}, 0)$ is by lemma (2.3) of dimension $n-2$.

To see that $\text{aff}(\mathcal{N}_{abc}, 0)$ is equal to $\mathcal{I}_{abc} := \mathcal{I}_{ab} \cap \mathcal{I}_{bc} \cap \mathcal{I}_{ac}$, we note once again that the affine hull of a set is the intersection of all the flats that contain that set, and $\{\mathcal{N}_{abc}, 0\}$ is a subset of $\{\mathcal{N}_{..}, 0\}$. The fact that the sets $\{\mathcal{I}_{..}\}$ are distinct (because the sets $\{\mathcal{H}_{..}\}$ are distinct) and $n-1$-dimensional then completes the argument. Therefore, in an identical manner to the case for $\mathcal{H}_{abc}^{\perp}$ in proposition (2.3), the 2-dimensional subspace of $\mathcal{R}^n$, $\mathcal{I}_{abc}^{\perp}$, is isomorphic to $\mathcal{R}^n/\mathcal{I}_{abc}$ under the correspondence $x \mapsto x + \mathcal{I}_{abc}$ for $x \in \mathcal{I}_{abc}^{\perp}$.

The fact that $\mathcal{I}_{..}$ has 1-dimensional intersection with $\mathcal{I}_{abc}^{\perp}$ follows from the fact that $\mathcal{I}_{..}^{\perp}$ is a 1-dimensional subset of $\mathcal{I}_{abc}^{\perp}$ and $\mathcal{I}_{..} + \mathcal{I}_{..}^{\perp} = \mathcal{R}^n$. (Recall that $\mathcal{I}_{abc}$ is a subset of $\mathcal{I}_{..}$ if and only if $\mathcal{I}_{..}^{\perp}$ is a subset of $\mathcal{I}_{abc}$.) Thus $\mathcal{I}_{..}$ is a hyperplane through the origin that determines two half-spaces in $\mathcal{I}_{abc}^{\perp}$. These half-spaces are equal to $\mathcal{Y}_{..} \cap \mathcal{I}_{abc}^{\perp}$ and $\mathcal{Z}_{..} \cap \mathcal{I}_{abc}^{\perp}$, and the former contains the set $\mathcal{I}_{..}^{\perp} \cap \mathcal{Y}_{..}$. Therefore, the set $\mathcal{I}_{abc}^{\perp}$ contains any candidate triple f_{ab}, f_{bc} and f_{ac} of vectors that satisfy

$$f_{ab} + f_{bc} = f_{ac}$$

as well as being perpendicular to the boundary of and pointing into their respective half-spaces $\mathcal{Y}_{ab}$, $\mathcal{Y}_{bc}$ and $\mathcal{Y}_{ac}$ as required for theorem (2.1). Using the same technique as in proposition (2.3), we will now find such a triple.

By proposition (2.3), there exists $r \in \operatorname{ri}\Delta \cap \mathcal{I}_{abc}$. Since $\Delta \cap \mathcal{I}_{abc} = \mathcal{N}_{abc}$, r lies in $\mathcal{N}_{abc}$, which by condition (NT) is equal to $\mathcal{N}_{ab} \cap \mathcal{N}_{bc}$. Then since $\mathcal{N}_{ab}$ and $\mathcal{N}_{bc}$ are, by condition (Div.(i)), distinct, let $p \in \operatorname{ri}\Delta \cap (\mathcal{N}_{ab} \backslash \mathcal{N}_{abc})$ and $q \in \operatorname{ri}\Delta \cap (\mathcal{N}_{bc} \backslash \mathcal{N}_{abc})$ such that $\operatorname{aff}(p,r) \neq \operatorname{aff}(q,r)$. Such a pair exist because firstly $\operatorname{ri}\Delta \cap \mathcal{N}_{..}$ is nonempty and of dimension $n-2$ by lemma (2.1), whilst $\mathcal{N}_{abc}$ is of dimension $n-3$. Secondly, $\operatorname{aff}(p,r) \subset \operatorname{aff}(p, \mathcal{N}_{abc}) = \operatorname{aff}(\mathcal{N}_{ab}) = \mathcal{H}_{ab}$ is not equal to $\operatorname{aff}(\mathcal{N}_{bc})$ which contains $\operatorname{aff}(q,r)$.

Now since $\operatorname{aff}(p,r)$ and $\operatorname{aff}(q,r)$ are distinct lines with the point r in common, p, q and r form an independent set of points, so that the flat $\operatorname{aff}(p,q,r)$ is two dimensional. Since all three points lie in Δ, this flat is contained in $\operatorname{aff}\Delta$, and moreover, its intersection with $\mathcal{I}_{abc}$ is the singleton $\{r\}$. This implies that $\operatorname{aff}(p,q,r)$ and $\mathcal{I}_{abc}$ are complementary flats in $\mathcal{R}^n$. In turn, the fact that $\mathcal{I}_{ac}$ contains elements that lie outside $\mathcal{I}_{abc}$ implies that the intersection of $\mathcal{I}_{ac}$ with $\operatorname{aff}(p,q,r)$ is one dimensional, contains r, and by condition (Div.(i)), is distinct from $\operatorname{aff}(p,r)$ and $\operatorname{aff}(q,r)$.

Next consider the orthogonal projection $\widetilde{P} : \mathcal{R}^n \longrightarrow \mathcal{I}_{abc}^{\perp}$. The kernel of $\widetilde{P}$ is equal to $\mathcal{I}_{abc}$ so that for any $z \in \mathcal{I}_{abc}^{\perp}$, $\widetilde{P}^{-1}(z)$ is equal to $z + \mathcal{I}_{abc}$: the translation of $\mathcal{I}_{abc}$ by z. Moreover, the fact that $\widetilde{P}$ is continuous and linear, means that $\widetilde{P}(\mathcal{N}_{..})$ is a convex and closed subset of $\mathcal{I}_{..} \cap \mathcal{I}_{abc}^{\perp}$, and since $\mathcal{N}_{..}$ contains $\mathcal{N}_{abc}$, the projection of $\mathcal{N}_{..}$ contains the origin. Clearly, because p lies outside $\mathcal{I}_{abc}$, $\widetilde{P}(p) \neq 0$, and the same is true of q. We claim that $\widetilde{P}(p)$ lies in $\mathcal{I}_{ab} \cap \mathcal{I}_{abc}^{\perp}$ and $\widetilde{P}(q)$ lies in $\mathcal{I}_{bc} \cap \mathcal{I}_{abc}^{\perp}$.

Claim 2.1. *$\widetilde{P}(p)$ lies in $\mathcal{I}_{ab} \cap \mathcal{I}_{abc}^{\perp}$ and $\widetilde{P}(q)$ lies in $\mathcal{I}_{bc} \cap \mathcal{I}_{abc}^{\perp}$.*

First note that $\mathcal{I}_{abc}$ is of dimension one less than $\mathcal{I}_{ab}$, and that both are subspaces, so that $\mathcal{I}_{abc}$ determines two half-spaces in $\mathcal{I}_{ab}$. (By taking an appropriate rotation, we see that they are homeomorphic to canonical subspaces of $\mathcal{R}^n$ that differ by one dimension.) Now let the open half-space to which p belongs be denoted by K. By proposition (2.2) K is a convex cone and $\mathcal{I}_{abc}$ is its (relative) boundary. Then for any element z of $\mathcal{I}_{abc}$, $p\frac{1}{2}z$ lies in K because K is convex. Then, by the fact that it is a cone and $2(1/2p + 1/2z) = p + z$ lies in K. Therefore, $p + \mathcal{I}_{abc}$ is a subset of $\mathcal{I}_{ab}$, and because $\widetilde{P}(p)$ equal to $(p + \mathcal{I}_{abc}) \cap \mathcal{I}_{abc}^{\perp}$, the proof of the claim is complete. An identical argument holds for q and $\mathcal{I}_{bc} \cap \mathcal{I}_{abc}^{\perp}$.

Now the fact that $\widetilde{P}$ is continuous and linear means that the image of $\operatorname{aff}(p,q,r)$ under $\widetilde{P}$ is a two dimensional subspace of (and hence equal to) $\mathcal{I}_{abc}^{\perp}$. Let $P : \operatorname{aff}(p,q,r) \longrightarrow \mathcal{I}_{abc}^{\perp}$ be the restriction of $\widetilde{P}$ to $\operatorname{aff}(p,q,r)$. Then P is a bounded linear isomorphism and as such, by the open mapping theorem, it maps open sets to open sets. It is therefore a homeomorphism between the two spaces, and moreover, P^{-1} is also a projection operator with identical properties to P, in that $P^{-1}(x)$ is equal to $(x + \mathcal{I}_{abc}) \cap \operatorname{aff}(p,q,r)$.

Let the intersection of sets such as $\operatorname{ri}\Delta$ with $\operatorname{aff}(p,q,r)$ be denoted by $\operatorname{ri}\Delta'$. By the above, $P(\operatorname{ri}\Delta')$ is open in $\mathcal{I}_{abc}^{\perp}$ and moreover as r is an element of $\operatorname{ri}\Delta'$, $0 \in P(\operatorname{ri}\Delta')$. This implies that we are free to choose a closed disc $\mathcal{D} \subset \mathcal{I}_{abc}^{\perp}$ of fixed radius ϵ around the origin, that is small enough to be contained in $\operatorname{ri}\Delta'$. That is, small enough so that $P^{-1}(\mathcal{D})$ is contained in $\operatorname{ri}\Delta'$. Let $\mathcal{S}$ be the circle that defines the boundary of this disc. For the purposes of the next claim we denote the image of any set X under P by $\widetilde{X}$.

Claim 2.2. *In $\mathcal{I}_{abc}^{\perp}$, the open half-space , $\widetilde{\mathcal{Y}_{..}}$ is equal to* $\mathrm{cone}\,\widetilde{\mathcal{B}_{..}} - \{0\}$*. Furthermore, for all $i, j, k, l \in \{a, b, c\}$ such that $i \neq j$ and $k \neq l$,*

$$\widetilde{\mathcal{Y}_{ij}} \cap \widetilde{\mathcal{Y}_{kl}} = \mathrm{cone}(\widetilde{\mathcal{B}_{ij}} \cap \widetilde{\mathcal{B}_{kl}}) - \{0\}.$$

First note that an identical argument to the one we made in claim (2.1) for $\mathcal{I}_{ab}$ shows that because $\mathcal{I}_{abc}$ is the relative boundary of the open convex cone $\mathcal{Y}_{..}$ in $\mathcal{R}^n$, the projection of $\mathcal{Y}'_{..}$, is equal to $\widetilde{\mathcal{Y}_{..}} = \mathcal{Y}_{..} \cap \mathcal{I}_{abc}^{\perp}$. Now the fact the $\mathcal{B}_{..}$ is a subset of $\mathcal{Y}_{..}$ implies that $\widetilde{\mathcal{B}_{..}}$ is a subset of $\widetilde{\mathcal{Y}_{..}}$. Indeed, proposition (2.2) implies that the same is true for each of the cones $\widetilde{\mathcal{Y}_{ij}} \cap \widetilde{\mathcal{Y}_{kl}}$ for $i, j, k, l \in \{a, b, c\}$ such that $i \neq j$ and $k \neq l$, and, with the exception of the cases where both $k = j$ and $l = i$, these cones are nonempty by distinctness of the sets $\mathcal{I}'_{..}$.

As $\widetilde{\mathcal{Y}_{..}}$ is an open half-space in $\mathcal{I}_{abc}^{\perp}$, and its boundary, $\mathcal{I}_{..}$ contains the origin, proposition (2.2) implies that $\widetilde{\mathcal{Y}_{..}} = \mathrm{cone}(\mathcal{S} \cap \mathcal{Y}_{..}) - \{0\}$. Now by the preceding paragraph $\mathcal{S} \cap \mathcal{Y}_{..}$ is a nonempty subset of $\mathcal{D}$, and so its pre-image under P is a nonempty subset of $\Delta' \cap \mathcal{Y}'_{..} \equiv \mathcal{B}_{..} \cap \mathrm{aff}(p, q, r)$. So therefore, $P(\mathcal{B}'_{..}) = \widetilde{\mathcal{B}_{..}}$ contains the set $\mathcal{S} \cap \mathcal{Y}_{..}$ and the first part of the claim follows. The second part follows by an identical argument.

The above claim will allow us to work exclusively on $\mathcal{I}_{abc}^{\perp}$. So in the knowledge that $x \in \widetilde{\mathcal{Y}_{..}}$ if and only if there exists $\lambda > 0$ such that $P^{-1}(\lambda x) \in \mathcal{B}_{..}$, for the remainder of this proof we avoid superfluous notation and write of projections to $\mathcal{I}_{abc}^{\perp}$ without explicit reference to the fact. For example $\mathcal{Y}_{..}$, $\mathcal{B}_{..}$ and $\mathcal{N}_{..}$, will refer to $\widetilde{\mathcal{Y}_{..}}$, $\widetilde{\mathcal{B}_{..}}$ and $\widetilde{\mathcal{N}_{..}}$ respectively.

Let $B = \{b_1, b_2\}$ be a pair of of orthonormal vectors in $\mathcal{I}_{abc}^{\perp}$ and let $C = \{c_1, \ldots, c_{n-2}\}$ be any orthonormal ordered basis for the $n-2$-dimensional subspace $\mathcal{I}_{abc}$ such that $\{B, C\}$, which defines an orthonormal basis for $\mathcal{R}^n$, listed in this particular order is equivalently oriented to the standard basis $\{\delta_1, \ldots, \delta_n\}$. That is, the matrix that consists of the vectors that define the basis we have chosen,

$$G = \begin{bmatrix} b_1, & b_2, & c_1, & \ldots, & c_{n-2} \end{bmatrix}$$

has positive determinant. Now define the function $R : [0, 2\pi] \times \mathcal{R}^n \longrightarrow \mathcal{R}^n$ such that for each $\theta \in [0, 2\pi]$, $R_\theta := R(\theta, \cdot)$ is defined to be the composition

$$G \circ \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 & \ldots \\ \sin(\theta) & \cos(\theta) & 0 & \ldots \\ 0 & 0 & 1 & \ddots \\ \vdots & \vdots & \ddots & \ddots \end{bmatrix} \circ G^*.$$

Then R defines a rotation on $\mathcal{I}_{abc}^{\perp}$, for it leaves vectors that are spanned by C unchanged. To see this, consider $R(\theta, c_1)$. Since $\{B, C\}$ defines an orthonormal basis, $G^*(c_1) = \delta_3$ as every row of G^* is the transpose of a vector in the basis. Then δ_3 is clearly left unchanged by the second matrix in the composition. Finally, since $G(\delta_3) = c_1$, we see that for any given value of θ, a vector in $\mathcal{I}_{abc}$ is unchanged by composition with $R(\theta,)$. For any given $x \in \mathcal{R}^n$, will also use the notation R_x as shorthand for the function $R(\cdot, x) : [0, 2\pi] \longrightarrow \mathcal{R}^n$; no confusion should arise because we will always use Latin subscripts for elements of $\mathcal{R}^n$ and Greek subscripts for elements of $[0, 2\pi]$.

Let p be an element of $\mathcal{N}_{ac} \cap \mathcal{S}$ such that for any $0 < \theta < \pi$, $R_p(\theta)$ is an element of $\mathcal{B}_{ac}$ and for all $\pi < \theta < 2\pi$, $R_p(\theta) \in \mathcal{W}_{ac}$. We can do this without loss of generality because if $y = -p$, then by claim (2.2), $R_y(\theta) = R_p(\theta + \pi)$ must lie in the opposite semi-circle to $R_p(\theta)$.

Now as both $\mathcal{N}_{ab} \cap \mathcal{S}$ and $\mathcal{N}_{bc} \cap \mathcal{S}$ are also nonempty, and by the same argument as in proposition (2.3), we know that there epist θ_{ab} and θ_{bc} in $]0, \pi[$ such that $R_p(\theta_{ab}) \in \mathcal{N}_{ab}$ and $R_p(\theta_{bc}) \in \mathcal{N}_{bc}$. We will now show that there are two types of arrangements for the hyperplanes that define preferences satisfying (NT).

First suppose that $\theta_{ab} < \theta_{bc}$, and that for any $0 \leqslant \zeta < \theta_{ab}$, $R_p(\zeta)$ lies in $\mathcal{B}_{ab}$. Then since θ_{ab} lies in $]0, \pi[$, the point $\theta_{ab} + \pi$ lies in $]\pi, 2\pi[$, so that for all $\theta_{ab} < \zeta < \pi$ we have $R_p(\zeta) \in \mathcal{W}_{ab} \cap \mathcal{B}_{ac} = \{b \succ_. a \succ_. c\}$. Then condition (NT) implies that for all such ζ, the vector $R_p(\zeta)$ lies in $\mathcal{B}_{bc}$, thereby contradicting the fact that θ_{bc} lies between 0 and π.

Instead, we maintain the assumption that $\theta_{ab} < \theta_{bc}$, and suppose that for all $0 \leqslant \zeta < \theta_{ab}$ we have $R_p(\zeta) \in \mathcal{W}_{ab}$. Then for all such ζ we have $R_p(\zeta) \in \mathcal{B}_{ba} \cap \mathcal{B}_{ac} \equiv \{b \succ_. a \succ_. c\}$, so that by (NT), we see that $R_p(\zeta) \in \mathcal{B}_{bc}$. Furthermore, as we rotate anti-clockwise through the threshold corresponding the value θ_{ab} strict preference for b over a changes to strict preference for a over b, so that for $\theta_{ab} < \zeta < \theta_{bc}$, the vector $R_p(\zeta)$ is an element of $\mathcal{B}_{ab} \cap \mathcal{B}_{bc} = \{a \succ_. b \succ_. c\}$ which is a subset of $\mathcal{B}_{ac}$ as required.

Indeed an iteration of the above argument shows no contradictions arise, and the following "table" describes preferences when $0 < \theta_{ab} < \theta_{bc} < \pi$:

$$R_p(\zeta) \in \begin{cases} \{b \succ_{.} a \gtrsim_{.} c\} & \Leftrightarrow \quad 0 \leqslant \zeta < \theta_{ab} \\ \{a \gtrsim_{.} b \succ_{.} c\} & \Leftrightarrow \quad \theta_{ab} \leqslant \zeta < \theta_{bc} \\ \{a \succ_{.} c \gtrsim_{.} b\} & \Leftrightarrow \quad \theta_{bc} \leqslant \zeta < \pi \\ \{c \gtrsim_{.} a \succ_{.} b\} & \Leftrightarrow \quad \pi \leqslant \zeta < \theta_{ab} + \pi \\ \{c \succ_{.} b \gtrsim_{.} a\} & \Leftrightarrow \quad \theta_{ab} + \pi \leqslant \zeta < \theta_{bc} + \pi \\ \{b \gtrsim_{.} c \succ_{.} a\} & \Leftrightarrow \quad \theta_{bc} + \pi \leqslant \zeta < 2\pi \end{cases} \tag{2.2}$$

We conclude therefore, that as we rotate anti-clockwise through the 6-tuple of thresholds, transitivity dictates that preferences permute the order of a single pair of elements of A, and as such these pairs must be neighbours under the order. To be concise, preferences permute through the following sequence as we rotate anti-clockwise: $bac \to abc \to acb \to cab \to cba \to bca$, before returning to bac.

An identical argument shows that if $0 < \theta_{bc} < \theta_{ab} < \pi$, then preferences permute through the opposite sequence as we rotate anti-clockwise: $acb \to abc \to bac \to bca \to cba \to cab$, and then back to acb.

Consider the case where $0 < \theta_{ab} < \theta_{bc} < \pi$. We claim that if we take $y := R_p(\theta_{ab})$, then $(y + \mathcal{I}_{bc}) \cap \mathcal{I}_{ac}$ is nonempty. An identical argument holds for the case $\theta_{ab} < \theta_{bc}$, where instead we take $y' = R_p(\theta_{bc})$ and show that $(y' + \mathcal{I}_{ab}) \cap \mathcal{I}_{ac}$ is nonempty.

Note that since $y := R_p(\theta_{ab})$ and $\theta_{ab} < \theta_{bc}$, from the above we know that

$y \in \mathcal{B}_{bc}$. Then since $\mathcal{I}_{bc}$ is the boundary of the half-space $\mathcal{Y}_{bc}$, $y + \mathcal{I}_{bc}$ is a subset of $\mathcal{Y}_{bc}$. Now since $\mathcal{I}_{bc}$ and $\mathcal{I}_{ac}$ are complementary subspaces of $\mathcal{I}^{\perp}_{abc}$ and $y + \mathcal{I}_{bc}$ is a coset of $\mathcal{I}_{bc}$, $y + \mathcal{I}_{bc}$ and $\mathcal{I}_{ac}$ are also complementary subspaces of $\mathcal{I}^{\perp}_{abc}$. Since $\mathcal{I}_{bc}$ and $\mathcal{I}_{ac}$ intersect only at the origin, $y + \mathcal{I}_{bc}$ and $\mathcal{I}_{ac}$ intersect at a unique vector other than the origin, which we denote by z.

On the one hand, since z is an element of $y + \mathcal{I}_{bc} \subset \mathcal{Y}_{bc}$, we know that for some $\lambda > 0$, $\lambda z \in \mathcal{B}_{bc} \cap \mathcal{S}$. On the other, z lies in $\mathcal{I}_{ac}$. This implies that either $\lambda z = R_p(0)$, or $\lambda z = R_p(\pi)$, both of which are elements of $\mathcal{N}_{ac}$. At λz therefore, preferences satisfy $b \succ_{.} c \sim_{.} a$, and the table of preferences (2.2) tells us that $0 \leqslant \zeta < \theta_{ab}$. Now since $\theta_{ab} \leqslant \pi$, we conclude that $\zeta = 0$ and $\lambda z = p$.

Now consider the vector $z - y$. The fact that z lies in $y + \mathcal{I}_{bc}$, together with the definition of this set implies that $z - y \in \mathcal{I}_{bc}$. Once again we need to establish the orientation of x. Now since $y = R_p(\theta_{ab})$, $-y$ is equal to $R_p(\theta_{ab} + \pi)$, so that by table (2.2) we see that $-y \in \{c \succ_{.} b \succsim_{.} a\}$. This implies that $-y$ is an element of $\mathcal{Y}_{ca}$, and since z is an element of the boundary, $\mathcal{I}_{ac}$, of $\mathcal{Y}_{ca}$, $\frac{1}{2}z + \frac{1}{2}(-y) \in \mathcal{Y}_{ca}$ by convexity and then by the fact that $\mathcal{Y}_{ca}$ is a cone, $2(\frac{1}{2}z + \frac{1}{2}(-y)) = z - y$ is in $\mathcal{Y}_{ca}$.

We conclude therefore that $x := z - y$ is an element of $\mathcal{I}_{bc} \cap \mathcal{Y}_{ca}$, and so for some $\mu > 0$, preferences at μx satisfy $b \sim_{\mu x} c \succ_{\mu x} a$, and so table (2.2) tells us that $\mu x = R_p(\theta_{bc} + \pi)$.

Next we show that the rotation of the vectors x, y and z in the same, anti-clockwise direction by $\frac{\pi}{2}$ lies in the intersection of the appropriate half-space with the orthogonal complement of the respective hyperplane to which they belong.

To this end, note that because $\theta_{ab} < \theta_{ab} + \frac{\pi}{2} < \theta_{ab} + \pi$, at $y' = R_p(\theta_{ab} + \frac{\pi}{2})$, by table (2.2), preferences satisfy $a \succ_{y'} b$, so that $y' \in \mathcal{Y}_{ab}$. Moreover, since $R_p(\theta_{ab} + \frac{\pi}{2}) = R_{\frac{\pi}{2}}(R_p(\theta_{ab})) = R(y, \frac{\pi}{2})$, we see that y' is indeed the rotation of y anti-clockwise by a right-angle onto the set $\mathcal{I}_{ab}^{\perp}$.

Similarly, since $\lambda z = p = R_p(0)$, and $0 < \frac{\pi}{2} < \pi$, preferences at $R(p, \frac{\pi}{2}) \in \mathcal{I}_{ac}^{\perp}$ satisfy $a \succ_{\cdot} c$. Therefore $z' := \frac{1}{\lambda} R(p, \frac{\pi}{2})$, which, by linearity of R in its first argument, is equal to $R(z, \frac{\pi}{2})$, is an element of $\mathcal{Y}_{ab}$.

In the same way, for the vector $\mu x = R_p(\theta_{bc} + \pi) \in \mathcal{I}_{bc}$, the fact that $\frac{3}{2}\pi < 2\pi$ implies that $\theta_{bc} + \frac{3\pi}{2} \pmod{2\pi}$ lies in $[0, \theta_{bc}) \cup (\theta_{bc} + \pi, 2\pi]$. Therefore, by an identical argument to the above, we see that the rotation under R of μx by $\frac{\pi}{2}$ anti-clockwise, is equal to $R_p(\theta_{bc} + \frac{3\pi}{2})$, an element of $\mathcal{Y}_{bc} \cap \mathcal{I}_{bc}^{\perp}$. Then letting $x' := \frac{1}{\mu} R_p(\theta_{bc} + \frac{3\pi}{2})$, we see that because

$$\frac{1}{\mu} R(R_p(\theta_{bc} + \pi), \pi/2)) = R(x, \pi/2),$$

and $x = z - y$, we have

$$x' = R(x, \pi/2) = R(z - y, \pi/2) = R(z, \pi/2) - R(y, \pi/2) = z' - y'.$$

Now if we define $f_{ab} = y'$, $f_{bc} = x'$ and $f_{ac} = z'$, we see that

$$f_{ab} + f_{bc} = f_{ac}$$

Then, defining $F'_{..}$ to be the linear functional on $\mathcal{R}^n$ defined by the inner product $\langle f_{..}, - \rangle$ on $\mathcal{R}^n$ and $F_{..}$ its restriction to the set Δ, we see that for all $p \in \Delta$

$$F_{ab}(p) + F_{bc}(p) = F_{ac}(p),$$

as required. ☐

Theorem (2.2) together with lemma(2.4) show that for $|A| \leqslant 3$ the pairwise conditions together with (NT) and Div.(i)) are sufficient for the existence of an expected utility function that is CUC across states.

In order to obtain the corresponding result for the case where $|A| \geqslant 4$, we make use of condition (Div.(ii)). This condition, is not necessary for the existence of an expected utility function, but, like condition (Div.(i)), it is needed in the sufficiency argument to ensure that the system of equations is consistent. In particular, (Div.(ii)) ensures that preferences such as those we observe in figure (2.4) are ruled out.

Lemma 2.5. *Let $|A| > 3$. If preferences satisfy conditions (Asy.), (C'ty), (WP), (NT), (p-SB), (Div.i) and (Div.ii), then for all distinct $a, b, c, d \in A$,*

$$\mathcal{I}_{ab}^{\perp} = C := \mathcal{I}_{abc}^{\perp} \cap \mathcal{I}_{abd}^{\perp}$$

Proof. Let a, b, c and d be four distinct elements of A. The fact that $\mathcal{I}_{ab}^{\perp}$ is

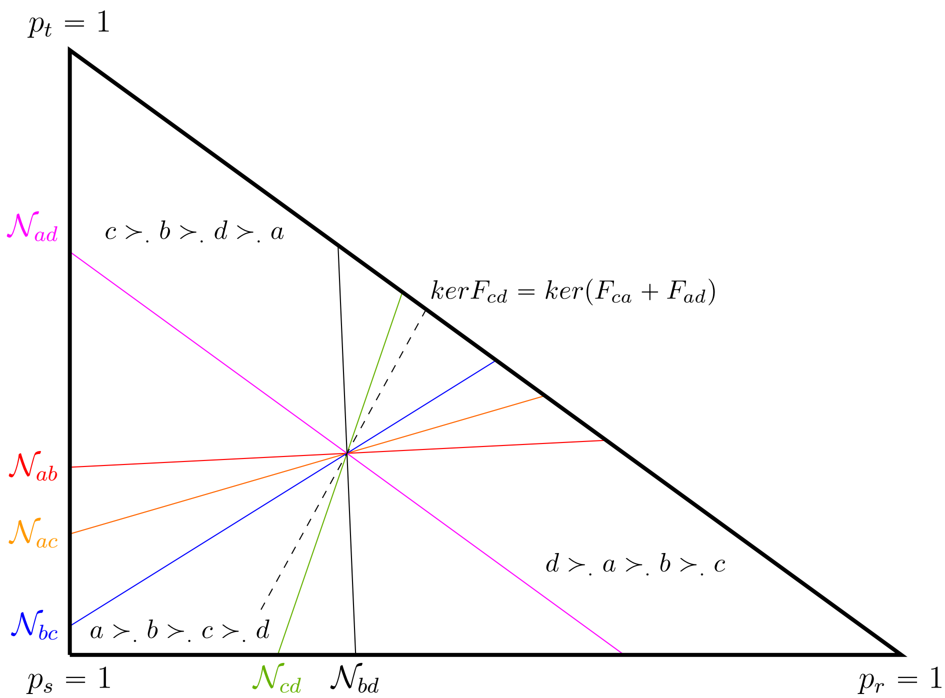


Figure 2.4: The preferences in this figure clearly fail to satisfy (Div.ii) for at the point $p = (\frac{1}{3}, \frac{1}{3})$, we have $a \sim_p b \sim_p c \sim_p d$, and there is no other point such that $a \sim_p b \sim_p c$. If the set $\mathcal{N}_{cd}$ (in green) happened to lie on the dashed line that defines $\ker(Fca + Fad)$, then there would be an expected utility representation. There is, however, no behavioural reason why the agent should have such preferences.

a subset of both $\mathcal{I}_{abc}^{\perp}$ and $\mathcal{I}_{abd}^{\perp}$ follows from the fact that, by construction, $\mathcal{I}_{abc}$ and $\mathcal{I}_{abd}$ are both subsets of $\mathcal{I}_{ab}$. For the reverse inclusion, since $\mathcal{I}_{ab}^{\perp}$ is a 1-dimensional subspace of C, and C is itself a subspace, we know that C is equal to $\mathcal{I}_{ab}^{\perp}$ if it is 1-dimensional. As $\mathcal{I}_{abc}^{\perp}$ and $\mathcal{I}_{abc}^{\perp}$ are both 2-dimensional, their intersection is at most 2-dimensional and C attains this dimension if and only if $\mathcal{I}_{abc}^{\perp} = \mathcal{I}_{abd}^{\perp} = C$.

We will suppose that this is true and seek a contradiction of condition (Div.ii). First note that since $\mathcal{I}_{abc}^{\perp} = \mathcal{I}_{abd}^{\perp}$ and both are 2-dimensional, their orthogonal complement $C^{\perp}$ is a subspace of dimension $n - 2$. Moreover, as $\mathcal{I}_{abc}$ and $\mathcal{I}_{abd}$

are both orthogonal to C, $C^{\perp}$ contains both of these $n-2$-dimensional subspaces of $\mathcal{R}^n$. We conclude, therefore, that $\mathcal{I}_{abc}$ and $\mathcal{I}_{abd}$ are equal.

This implies that $\mathcal{N}_{abc} = \Delta \cap \mathcal{I}_{abc} = \Delta \cap \mathcal{I}_{abd} = \mathcal{N}_{abd}$. That is $a \sim_p b \sim_p c$ holds if and only if $a \sim_p b \sim_p d$. By condition (NT) we conclude that the set $\{p : a \nsim_p b \sim_p c \sim_p d\}$ is empty, a contradiction of condition (Div.ii). $\square$

Theorem 2.3. *The following two statements are equivalent.*

1) *Context preferences $\{(A, \succ_p) : p \in \Delta\}$ satisfy conditions (Asy.), (C'ty), (WP), (p-SB), (NT) and (Div.).*

2) *Context preferences have an expected* preference *representation $F : A \times A \times \Delta \to \mathcal{R}$ with the strong groupoid property. The vectors $\{f_{ij} \in \mathcal{R}^n : i, j \in A\}$ that characterize F satisfy both the following:*

   i) *for every list (a, b, c) of 3 distinct elements in A, the closed cord between f_{ab} and f_{bc} lies outside the closure of the negative orthant;*

   ii) *for every list (a, b, c, d) of 4 distinct elements in A, f_{ab} and f_{bc} and f_{cd} are linearly independent.*

Note that when $|S| = 2$, there are no preferences satisfying the conditions, and so the theorem is trivially satisfied.

Proof. We prove that (1)$\Rightarrow$ (2) by showing that (1) and not (2) fails to hold. First we show that (1) implies the existence of an expected preference representation with the strong groupoid property. Thus if not (2) is to hold, then it must be that not (2.i) or not (2.ii). Once this is shown to be false, the only remaining possibility is the required one.

We then argue that (2)$\Rightarrow$(Div.), for (NT) is implied by the strong groupoid property and via theorem (2.1) the existence of an expected preference representation implies (Asy.), (C'ty), (WP) and (p-SB). So we show that a strong groupoid expected preference representation that satisfies (2.i) must satisfy (Div.i). Finally, we show that a (2) is also sufficient for (Div.ii).

(1)$\Rightarrow$ (2): Let a, b be a distinct pair in A, and take f_{ab} to be the vector in $\mathcal{I}_{ab}^{\perp} \cap \mathcal{Y}_{ab}$ that has length one. Now for all $x \in A$, a, b, x is either a triple of distinct elements of A, or $x \in \{a, b\}$. Now when they form a triple, lemma (2.4) implies that we may find unique vectors $f_{bx} \in \mathcal{I}_{bx}^{\perp} \cap \mathcal{Y}_{bx}$ and $f_{ax} \in \mathcal{I}_{ax}^{\perp} \cap \mathcal{Y}_{ax}$

$$f_{ab} + f_{bx} = f_{ax}, \text{ or, equivalently } f_{bx} = f_{ax} - f_{ab}. \tag{2.3}$$

Whenever x is equal to b, then we can still choose the vector f_{bx} to solve the equation $f_{ab} + f_{bb} = f_{ab}$, so that f_{bb} is the unique element of $\mathcal{I}_{bb}^{\perp} = (\mathcal{R}^n)^{\perp} = \{0\}$. If $x = a$, then the same equation yields $f_{ab} + f_{ba} = f_{aa}$, and so if we choose $f_{aa} = 0$, then $f_{ba} = -f_{ab} \in \mathcal{I}_{ab}^{\perp} \cap \mathcal{Y}_{ba}$.

So far, we have identified a collection of vectors $(f_{ij}, (i, j) \in \{a, b\} \times A)$. This collection not only satisfies the groupoid property, but also each member of the collection defines a linear operator $F_{yx} := \langle f_{yx}, \rangle$ such that $F_{yx}(p) > 0$ if and only if $y \succ_p x$.

Now for all $x \neq a, b$, let $f_{xb} := -f_{bx} \in \mathcal{I}_{bx}^{\perp} \cap \mathcal{Y}_{xb}$ and $f_{xa} := -f_{ax} \in \mathcal{I}_{ax}^{\perp} \cap \mathcal{Y}_{xa}$.

Then for each x, the equation $f_{ab} + f_{bx} = f_{ax}$ is equivalent to

$$-f_{ax} = -f_{bx} - f_{ab} \quad \Leftrightarrow \quad f_{xb} + f_{ba} = f_{xa}\,;$$

and hence in the latter equation, by taking f_{ba} to the right-hand-side and using the fact that f_{ab} is the additive inverse of f_{ba}, we obtain

$$f_{xa} - f_{ba} = f_{xb} \quad \Leftrightarrow \quad f_{xa} + f_{ab} = f_{xb}\,;$$

and then multiplying the latter equation by -1, and once again substituting in each of the additive inverses,

$$-f_{xa} - f_{ab} = -f_{xb} \quad \Leftrightarrow \quad f_{ba} + f_{ax} = f_{bx}\,;$$

then once again in the latter equation, taking f_{ax} to the other side and using the fact that f_{ax} is the additive inverse of f_{ax}, we obtain

$$-f_{ax} + f_{bx} = f_{ba} \quad \Leftrightarrow \quad f_{bx} + f_{xa} = f_{ba}\,;$$

and finally, multiplying through be -1, and repeating the penultimate step, we obtain

$$-f_{bx} - f_{xa} = -f_{ba} \quad \Leftrightarrow \quad f_{ax} + f_{xb} = f_{ab}.$$

To summarize, so far we have identified a collection of vectors f_{ij} for $(i, j) \in (\{a, b\} \times A) \cup (A \times \{a, b\})$. This collection not only satisfies the groupoid property, but also each member of the collection defines a linear operator $F_{yx} := \langle f_{yx}, \rangle$ such that $F_{yx}(p) > 0$ if and only if $y \succ_p x$.

Now for each pair $c, d \in A$, take the two corresponding vectors f_{ac} and f_{ad}, that we have identified above, and define f_{cd} to be the element that satisfies $f_{cd} = f_{ad} - f_{ac} = f_{ca} + f_{ad}$. First note that $-f_{cd} = -(f_{ca} + f_{ad}) = f_{da} + f_{ca} \equiv f_{dc}$. Then since $f_{ad} \in \mathcal{I}_{ad}^{\perp}$ and $f_{ac} \in \mathcal{I}_{ac}^{\perp}$, f_{cd} lies in the direct sum $\mathcal{I}_{ad}^{\perp} \oplus \mathcal{I}_{ac}^{\perp}$, which, by the proof of lemma (2.4), is equal to $\mathcal{I}_{acd}^{\perp}$. So if f_{cd} is also an element of $\mathcal{I}_{bcd}^{\perp}$, then f_{cd} lies in $\mathcal{I}_{acd}^{\perp} \cap \mathcal{I}_{bcd}^{\perp}$ which, by lemma (2.5), is equal to $\mathcal{I}_{cd}^{\perp}$. We now show that this is true.

Consider the vectors f_{bc} and f_{bd}, defined in equation (2.3) immediately above. Any linear combination of these two vectors lies in $\mathcal{I}_{bc}^{\perp} \oplus \mathcal{I}_{bc}^{\perp} = \mathcal{I}_{bcd}^{\perp}$. Moreover, by equation (2.3) we see that,

$$\begin{aligned} f_{bd} - f_{bc} &= (f_{ad} - f_{ab}) - (f_{ac} - f_{ab}) \\ &= f_{ad} - f_{ac}, \end{aligned}$$

which is equal to f_{cd}, and therefore we conclude that $f_{cd} \in \mathcal{I}_{cd}^{\perp}$.

Note that the immediately preceding argument holds for $c = d$, and in this case we see that $f_{ac} = f_{ad}$, so that $f_{cd} \in \mathcal{I}_{cc}^{\perp} = \{0\}$ whenever $c = d$. It remains to be shown that f_{cd} is an element of $\mathcal{Y}_{cd}$ whenever c and d are distinct.

First suppose that $f_{cd} = 0$. This implies that $f_{ac} = f_{ad}$. In turn, this implies that $\mathcal{I}_{ac} = \mathcal{I}_{ad}$, which violates condition (Div.i). Now suppose that f_{cd} lies in $\mathcal{Z}_{cd} \equiv \mathcal{Y}_{dc}$. Now if we define $F_{cd} := \langle f_{cd}, \cdot \rangle$. Then for all $p \in \mathcal{W}_{cd} \equiv \mathcal{B}_{dc}$, we have $F_{cd}(p) > 0$, and since

$$F_{cd}(\cdot) = \langle f_{ad} - f_{ac}, \cdot \rangle = F_{ad}(\cdot) - F_{ac}(\cdot),$$

we conclude that for all such p, we have $F_{ad}(p) > F_{ac}(p)$.

Now consider the set $\{d \succ_. a \succ_. c\}$. This is clearly a subset of $\{d \succ_. c\} \equiv \mathcal{B}_{dc}$ and moreover, by condition (Div. i) it is nonempty. Thus for any p in this set, by construction, $F_{da}(p)$ and $F_{ac}(p)$ are both positive. Moreover, as $F_{ad} = -F_{da}$, we see that $F_{ad}(p) < 0$ and $F_{ac}(p) > 0$ for any $p \in \{d \succ_. a \succ_. c\} \subset \mathcal{B}_{dc}$, a contradiction of our conclusion in the preceding paragraph. Therefore, $f_{cd} \in \mathcal{I}_{cd}^{\perp} \cap \mathcal{Y}_{cd}$ as required.

We now show that for any c, d, d' and $e \in A$, $f_{cd} + f_{d'e} = f_{ce}$ is equivalent to $d = d'$. First note that

$$\begin{aligned} f_{cd} + f_{d'e} &= f_{ca} + (f_{ad} + f_{d'a}) + f_{ae} \\ &= f_{ca} + f_{ae} + (f_{ad} + f_{d'a}) \\ &= f_{ce} + (f_{ad} + f_{d'a}) \end{aligned}$$

so that $f_{cd} + f_{d'e} = f_{ce}$ if and only if the term in brackets is equal to 0. On the one hand if $d = d'$ then $f_{ad} + f_{da} = f_{aa} = 0$. On the other hand, if $f_{ad} + f_{d'a} = 0$, then $f_{ad} = -f_{d'a} = f_{ad'}$, so that the sets $\mathcal{I}_{ad}$ and $\mathcal{I}_{ad'}$ are equal. If $d \neq d'$, then, as we have seen the proof of lemma (2.2), we obtain a contradiction of condition (Div.i).

The above shows that there exists an expected preference representation $F : (A \times A) \times \Delta \to \mathcal{R}$. Moreover, since $F_{ii} = 0$ for all $i \in A$, the quotient set $G = \{F_{ij} : i, j \in A\}/\{F_{ij} : i = j\}$ is closed under the law of composition that is defined for F_{ij} and F_{kl} such that $k = l$. This set is in fact a group under this law of composition, for it has a unique identity element 0, it contains elements that have unique additive inverse in G, $F_{ij} = -F_{ji}$, and is clearly associative.

Finally, to see that the expected preference representation $F : (A \times A) \times \Delta \to \mathcal{R}$ is unique up to a positive scalar multiple, recall that the proof of theorem (2.1) shows that an expected preference representation is unique up to multiplication by a positive scalar. As the proof here is identical, we only need to show that the groupoid property is preserved under multiplication by a scalar.

Note that we chose f_{ab} to have norm one, and given this choice the family of vectors $\{f_{ij}, i, j \in A\}$ that we have constructed is, by lemma (2.4) unique. By the same lemma, it is also clear that for any other choice of positive norm, say $\theta > 0$,

$$\theta f_{ab} = \theta f_{ax} + \theta f x b$$

for all $x \in A$. Furthermore, our construction was such that for every $c, d \in A$, $f_{cd} = f_{ad} - f_{ac}$, an equality which is also preserved under multiplication by θ.

To complete the proof that statement (1) of our theorem implies (2), we fix a representation F of preferences that satisfies the strong groupoid property. We will first show that given such a representation, the negation of (2.i) implies the negation of condition (Div.i): a contradiction of our assumption that the latter is true.

If (2.i) fails to hold, then there exists a triple $\{a, b, c\}$ of distinct elements in A such that for some $0 \leqslant \lambda \leqslant 1$, the vector

$$f^{\lambda} := \lambda f_{ab} + (1 - \lambda) f_{bc} \leqslant 0,$$

where this inequality means that for every element f^{λ}_s of f^{λ}, $f^{\lambda}_s \leqslant 0$. Now take $p \in \mathcal{B}_{ab} \cap \mathcal{B}_{bc}$. Now since $p \in \Delta$, every element p_s of p satisfies $p_s \geqslant 0$, the inner product of f^{λ} and p, $\langle f^{\lambda}, p \rangle = \sum_s f^{\lambda}_s p_s$ must be non-positive.

Now by bi-linearity of the inner product,

$$\langle f^{\lambda}, p \rangle = \lambda \langle f_{ab}, p \rangle + (1 - \lambda) \langle f_{bc}, p \rangle \tag{2.4}$$

$$= \lambda F_{ab}(p) + (1 - \lambda) F_{bc}(p), \tag{2.5}$$

so that at least one of $F_{ab}(p)$ and $F_{bc}(p)$ is non-positive. However, since $p \in \mathcal{B}_{ab} \cap \mathcal{B}_{bc}$, this contradicts the fact that F is a representation.

To prove statement (2.ii), we suppose that our representation F satisfies (2.i) but not (2.ii) and seek a contradiction of (Div.(ii)).

Suppose that f_{cd} lies in the linear hull of f_{ab} and f_{bc}, $\mathrm{Lin}(f_{ab}, f_{bc}) = \mathcal{I}^{\perp}_{abc}$. Then since f_{cd} is also an element of both $\mathrm{Lin}(f_{bc}, f_{cd}) = \mathcal{I}^{\perp}_{bcd}$ and $\mathrm{Lin}(f_{ac}, f_{cd}) = \mathcal{I}^{\perp}_{acd}$, we conclude that

$$f_{cd} \in \mathcal{I}_{abc}^{\perp} \cap \mathcal{I}_{bcd}^{\perp} \cap \mathcal{I}_{acd}^{\perp},$$

which, by lemma (2.5), is equal to $\mathcal{I}_{bc}^{\perp} \cap \mathcal{I}_{cd}^{\perp} \cap \mathcal{I}_{ac}^{\perp}$. Now since statement (2) implies that $\mathcal{B}_{cd}$ and $\mathcal{W}_{cd}$ are both nonempty, and so $f_{cd} \neq 0$. Therefore, by proposition (2.1), $\mathcal{I}_{..}$ is of dimension $n-1$ and by lemma (2.1) equal to the linear hull of $\mathcal{N}_{..}$. Thus, each of the subspaces $\mathcal{I}_{..}^{\perp}$ are one dimensional and equal to $\mathcal{I}_{cd}^{\perp}$. This implies that $\mathcal{I}_{cd} = \mathcal{I}_{bc} = \mathcal{I}_{ac}$. Therefore, for all $p \in \mathcal{N}_{..} \equiv \mathcal{I}_{..} \cap \Delta$ we have $d \sim_p c \sim_p b$ and $a \sim_p c$. So that, via condition (NT), we conclude that $a \sim_p b \sim_p c \sim_p d$ for all such p: a violation of (Div.(ii)).

(2)⇒ (1): In the opposite direction, suppose that (2) holds and consider the following linear programming problem which is based on the one found in Gilboa and Schmeidler (1999) and which we denote by (P^*).

$$\begin{aligned} &\min_{p \in \Delta} \langle 0, p \rangle \\ s.t. \quad & F_{ab}(p) = \langle f_{ab}, p \rangle \geqslant 1 \\ & F_{bc}(p) = \langle f_{bc}, p \rangle \geqslant 1. \end{aligned}$$

The fact that F is a representation of preferences implies that if this problem is feasible, then for any p in the feasible set, $a \succ_p b \succ_p c$ and (Div.(i)) is satisfied. Clearly, if a problem is soluble (has an optimal vector) it is feasible, and Theorem 4.3.4 of Webster p.161 states that (P^*) is soluble if its primal, (P), is soluble, so we now show that (P) is soluble.

Rewriting (P) into standard form, so that we minimize y in the closure of the positive orthant $\bar{\mathcal{R}}^n_+$, means we include the following additional constraints to ensure that we are choosing $y \in \Delta$:

$$\langle \mathbf{1}, y \rangle \geqslant 1 \quad \text{and} \quad \langle -\mathbf{1}, y \rangle \geqslant -1,$$

where $\mathbf{1}$ is a vector of ones in $\mathcal{R}^n$. In standard form therefore, (P) may be written as

$$\begin{aligned} &\min_{y \in \bar{\mathcal{R}}^n_+} \langle \mathbf{b}, y \rangle \\ &s.t. \quad \mathbf{A}(y) \geqslant \mathbf{c}, \end{aligned}$$

where $\mathbf{b}$ is the zero-vector in $\mathcal{R}^n$, and

$$\mathbf{A} = \begin{bmatrix} f^*_{ab} \\ f^*_{bc} \\ \mathbf{1}^* \\ -\mathbf{1}^* \end{bmatrix}_{(4\times n)} \qquad \mathbf{c} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \end{bmatrix}$$

Therefore, primal problem, (P), takes the form:

$$\begin{aligned} &\max_{x \in \bar{\mathcal{R}}^4_+} \langle \mathbf{c}, x \rangle = x_1 + x_2 + x_3 - x_4 \\ &s.t. \quad \mathbf{A}^*(x) = x_1 f_{ab} + x_2 f_{bc} + x_3 \mathbf{1} - x_4 \mathbf{1} \leqslant \mathbf{0}, \end{aligned}$$

Now since the objective function in (P) is decreasing in x_4, and $\mathbf{0} \in \mathcal{R}^4$ is feasible, any solution must set $x_4 = 0$. Now for any $\mathbf{x}'$ such that either x'_1 or

x_2' is positive, let $z = 1/(x_1' + x_2')$. Then because statement (2) holds, so does

$$x_1' z f_{ab} + x_2' z f_{bc} \nleqslant \mathbf{0}$$

as it is a convex combination of f_{ab} and f_{bc}. Moreover, this inequality is invariant under multiplication by positive constants, so we conclude that for all $\mathbf{x}$ such that either x_1 or x_2 is positive and $x_3, x_4 = 0$, are infeasible. Then because $\mathbf{1} = (1, \ldots, 1)^* \in \mathcal{R}^n$, for any $\mathbf{x}$ with $x_4 = 0$ that is not the zero vector,

$$x_1 f_{ab} + x_2 f_{bc} + x_3 \mathbf{1} \nleqslant \mathbf{0}.$$

This implies that $\mathbf{0}$ is the solution of (P).

Finally, the fact that (2) (and (2.ii) in particular) implies (Div.ii) is then deduced by the following brief argument. First, (2.ii) implies that for any distinct a, b, c and d in A, $f_{ab} \notin \text{span}\{f_{bc}, f_{cd}\} = \mathcal{I}_{bcd}^{\perp}$. This in turn implies $\mathcal{I}_{ab}^{\perp} \not\subset \mathcal{I}_{bcd}^{\perp}$, so that $\mathcal{I}_{bcd} \not\subset \mathcal{I}_{ab}$. Then, the fact that (2.i) implies (Div.i) means that for any set, such as Δ, that is transverse to both $\mathcal{I}_{ab}$ and $\mathcal{I}_{bcd}$, $\mathcal{I}_{bcd} \cap \Delta \equiv \mathcal{N}_{bcd}$ is not a subset of $\mathcal{N}_{ab}$. That is, (2.ii), by arguments implies the existence of $p \in \Delta$ such that $b \sim_p c \sim_p \sim_p d$ and $a \nsim b$.

□

Theorem 2.4. *The following two statements are equivalent.*

1) *Context preferences $\{(A, \succ_p) : p \in \Delta\}$ satisfy conditions (Asy.), (C'ty), (WP), (p-SB), (NT) and (Div.).*

2) *Context preferences have an CUC expected context utility representation $U : A \times \Delta \to \mathcal{R}$. The vectors $\{u_i \in \mathcal{R}^n : i \in A\}$ that characterize U*

satisfy both the following:

i) for every list (a, b, c) of 3 distinct elements in A, the closed cord between $u_a - u_b$ and $u_b - u_c$ lies outside the closure of the negative orthant;

ii) for every list (a, b, c, d) of 4 distinct elements in A, $u_a - u_b$, $u_b - u_c$ and $u_c - u_d$ are linearly independent.

Proof. This is a direct consequence of theorem (2.3) and theorem (2.2). □

2.4 Distinguishing the present model from that of Gilboa and Schmeidler (2003)

Proposition 2.4. *Suppose preferences $\{(A, \succ_p) : p \in \Delta\}$ satisfy (Asy.), (WP), (C'ty), (p-SB) and (NT). Then the following two statements are equivalent.*

(1) (Div.GS) *For every list (a, b, c, d) of distinct elements of A, there exists $p \in \Delta$ such that $a \succ_p b \succ_p c \succ_p d$. If $|A| < 4$, then for any strict ordering of the elements of A, there exists $p \in \Delta$ such that $\succ_p$ is that ordering.*

(2) (Div.(iii)) *For every list (a, b, c, d) of distinct elements of A, there exists $p, q \in \Delta$ such that*

$$d \succ_p a \sim_p b \sim_p c \quad \wedge \quad c \sim_q b \sim_q a \succ_q d$$

If $|A| = 3$, the corresponding statement holds in the absence of c, and if $|A| = 2$, then it holds in the absence of both b and c.

Moreover, if preferences also satisfy (Div.(iii)), then the dimension of $\mathcal{N}_{abcd} := \mathcal{N}_{ab} \cap \mathcal{N}_{bc} \cap \mathcal{N}_{cd}$ *is a convex set of dimension greater than or equal to* $n-4$.

Proof. Recall that our background assumption on A is that $|A| \geqslant 2$. For $|A| = 2$, the statements are immediately equivalent. For $|A| \geqslant 3$, the proof that (Div.(iii)) implies (Div.GS) is simple extension to the proof for $|A| = 3$. The latter proof proceeds as follows. Note that ($Div.(iii)$) states that there exists p_i, q_i, $i = 1, 2, 3$ in Δ such that

$$\begin{array}{lll} d \succ_{p_1} a \sim_{p_1} b & ; \quad b \sim_{q_1} a \succ_{q_1} d & ; \\ a \succ_{p_2} d \sim_{p_2} b & ; \quad b \sim_{q_2} d \succ_{q_2} a & ; \\ b \succ_{p_3} a \sim_{p_3} d & ; \quad d \sim_{q_3} a \succ_{q_3} b & . \end{array}$$

Now by condition (C'ty), there exists an open neighbourhood O of p_1 in Δ, such that for all $r \in O$, $d \succ_r a$ and $d \succ_r b$. Now since $a \sim_{p_1} b$ and $a \succ_{q_3} b$, by condition (p-SB), we know that for $0 < \lambda < 1$ sufficiently close to 1, $r' := p_1 \lambda q_3 \in O$, and $d \succ_{r'} a \succ_{r'} b$. Similarly, since $b \succ_{p_3} a$, there exists a convex combination r'' of p_3 and p_1 such that $r'' \in O$, and $d \succ_{r''} b \succ_{r''} a$.

A further four iterations of this argument (if necessary using (NT)) show that each of the 3! strict orders implied by (Div.GS) are implied by (Div.(iii)). For the converse we will use lemma (2.2). For the case where $|A| = 3$, the proof of lemma (2.2) shows that $\mathcal{N}_{bc}$ has nonempty intersection with both $\mathcal{B}_{ab}$ and $\mathcal{W}_{ab}$. That is, there exists $p, q \in \Delta$ such that

$$a \succ_p b \sim_p c \quad \wedge \quad b \sim_q c \succ_q a.$$

Now in that proof, the choice of a, b and c was arbitrary, thus (Div.GS) implies (Div.(iii)) for the case where $|A| = 3$.

For the case where $|A| = 4$, we note that (Div.GS) implies that for every permutation π of $\{a, b, c\}$, there exists $p^\pi, q^\pi \in \Delta$ such that

$$d \succ_{p^\pi} \pi(a) \succ_{p^\pi} \pi(b) \succ_{p^\pi} \pi(c) \quad \wedge \quad \pi(a) \succ_{q^\pi} \pi(b) \succ_{q^\pi} \pi(c) \succ_{q^\pi} d$$

So consider the set $\mathcal{B}_d(a, b, c) := \mathcal{B}_{da} \cap \mathcal{B}_{db} \cap \mathcal{B}_{dc}$ (and $\mathcal{W}_d(a, b, c)$ defined similarly). $\mathcal{B}_d(a, b, c)$ is the finite intersection of convex sets that are open in Δ, therefore it is convex and open in Δ, the same is true of $\mathcal{W}_d(a, b, c)$. Moreover, for each π, $p^\pi \in \mathcal{B}_d(a, b, c)$ and $q^\pi \in \mathcal{W}_d(a, b, c)$. The analogue of lemma (2.2) holds, therefore, in the topology of $\mathcal{B}_d(a, b, c)$ and $\mathcal{W}_d(a, b, c)$ respectively.

That is, the set $\mathcal{B}_d(a, b, c) \cap \operatorname{ri} \mathcal{N}_{abc}$ is nonempty and of dimension greater than or equal to $n-3$, and the same is true of $\mathcal{W}_d(a, b, c) \cap \operatorname{ri} \mathcal{N}_{abc}$. Since our choice of d was arbitrary, this completes the proof that (Div.GS) and (Div.(iii)) are equivalent.

For the last part of the proposition, we note that since $\mathcal{B}_d(a, b, c) \cap \operatorname{ri} \mathcal{N}_{abc}$ is of the same dimension as $\mathcal{N}_{abc}$, so is $\mathcal{B}_{da} \cap \operatorname{ri} \mathcal{N}_{abc}$ and $\mathcal{W}_{da} \cap \operatorname{ri} \mathcal{N}_{abc}$, and so $\mathcal{N}_{da} \cap \operatorname{ri} \mathcal{N}_{abc}$, by the same argument as the one we made in showing that $\dim \mathcal{N}_{abc} \geqslant n-3$ in lemma (2.2), disconnects the $n-3$-dimensional manifold $\operatorname{ri} \mathcal{N}_{abc}$, and is therefore of dimension greater than or equal to $n-4$. $\square$

The next propositions show that set of possible preferences that satisfy the pairwise conditions together with (NT) and (Div.GS) is empty for all $|A| \geqslant 3$

whenever $|S| = 2$ and it is empty for all $|A| \geqslant 4$ whenever $|S| = 3$. There is a related but somewhat more vague discussion of these facts in conclusion of Ashkenazi and Lehrer (2001).

Proposition 2.5. *(i) If $|S| = 2$, then for all $|A| \geqslant 3$, there are no preferences $\{(A, \succ_p) : p \in \Delta\}$ satisfying the pairwise conditions together with both (NT) and (Div.GS).*
(ii) If $|S| = 3$, then for all $|A| \geqslant 4$, there are no preferences satisfying the same conditions as in (i).
(iii) If $|S| \geqslant 4$, then for all $|A| \geqslant 4$, the set $\mathcal{N}_{abcd}$ is a convex set of dimension $n - 4$ whenever preferences satisfy the same conditions as in (i)

Proof. First consider the case where $|S| = 2$. In this case, Δ is of dimension 1, and by proposition (2.1) the sets $\mathcal{N}_{..}$ are singletons that are equal to their affine hull, and separate the sets $\mathcal{B}_{..}$ and $\mathcal{W}_{..}$. By proposition (2.4), (Div.GS) is equivalent to (Div.(iii)). So if there were preferences satisfying (Div.(iii)), then for $|A| \geqslant 3$, for all lists (a, b, c) of distinct elements of A, there exist two elements $p, q \in \Delta$ such that $a \succ_p b \sim_p c$ and $b \sim_p c \succ_p a$. This implies that $\mathcal{N}_{bc}$ is of cardinality greater than one, a contradiction.

For the case where $|S| = 3$, we again appeal to proposition (2.4), there we saw that the sets $\mathcal{B}_d(a, b, c) \cap \mathcal{N}_{abc}$ and $\mathcal{W}_d(a, b, c) \cap \mathcal{N}_{abc}$ are both nonempty. Now that fact that $\mathcal{N}_{abc}$ is the intersection of three convex sets, implies that it is itself convex and conditions (Asy.) and (C'ty) imply that the sets $\mathcal{B}_d(a, b, c)$ and $\mathcal{W}_d(a, b, c)$ are disjoint and open, so for any $p \in \mathcal{B}_d(a, b, c) \cap \mathcal{N}_{abc}$ and $q \in \mathcal{W}_d(a, b, c) \cap \mathcal{N}_{abc}$, $\text{conv}(p, q) \subset \mathcal{N}_{abc}$ is a one-dimensional set. Thus $\mathcal{N}_{abc}$ is of dimension at least equal to 1. However by lemma (2.3), we know that $\mathcal{N}_{abc}$

is of dimension $n - 3 = 0$, once more a contradiction.

For the case where $|S| = n \geqslant 4$, we note that by proposition (2.4), the dimension of $\mathcal{N}_{abcd} = N_{ad} \cap \mathcal{N}_{abc}$ is at least $n - 4$ and therefore nonempty for all . in Therefore, the sets $\mathcal{H}_{da} = \operatorname{aff} \mathcal{N}_{da}$ and $\mathcal{H}_{abc} = \operatorname{aff} \mathcal{N}_{abc}$ are nonempty subspaces in $\mathcal{R}^n$ that are of dimension $n - 2$ and $n - 3$ respectively. Then since $\mathcal{H}_{abc}$ is not a subset of $\mathcal{H}_{da}$, their direct sum is the whole of aff Δ, so that by theorem (1.3.8) of Webster, the dimension of their intersection is equal to $n - 4$. $\square$

2.5 Discussion of results

Firstly, we have improved the theorem of [GS03] to the extent that it now applies to a significantly larger class of preferences. In particular, our diversity condition is not so strong as to rule out all possible context preferences when the number of states is equal to 3 and the number of alternatives is greater than or equal to 4. That is, for $|S| = 3$ and any $|A| \geqslant 4$, the theorem of [GS03] is only true as a result of there being no preferences satisfying the conditions they impose; whereas in our model, the theorem is meaningful in such circumstances as the set of possible preferences that satisfy our diversity axiom is nonempty.

Even in dimensions higher than 3, our model allows for a substantially larger class of preferences, indeed, given the other conditions, the set of preferences satisfying (Div.GS) is of measure zero in the set preferences satisfying (Div.). Nevertheless, when the number of states is two and the number of alternatives is greater than 3, then as theorem (2.4) makes clear, there are no preferences satisfying the conditions in our model either. Note that the diversity condition may be restated so that the main theorem holds for $|S| \geqslant 3$ and, provided $|A| = 2$, $|S| = 2$ also. This would then be accompanied by a generic non-existence result for the case where $|A| \geqslant 3$ and $|S| = 2$. Indeed as Ashkenazi and Lehrer suggest, there appears to be no remotely intuitive condition on preferences that will extend the model to include the case where there are only two states.

Due to the fact that our model allows us to meaningfully consider the 2-dimensional simplex as our context space, we are able to present diagrams of possible arrangements of the sets, $\{\mathcal{N}_{ab} : a, b \in A\}$, to which context preferences give rise. Below, figure (2.5) presents an example of preferences that satisfy the conditions of the main theorem of this chapter when the matrix W with rows equal to the (transpose of the) vectors $\{u_a \in \mathcal{R}^n : a \in A\}$ for the case where $|S| = 3$ and $|A| = 4$. That is

$$W = \begin{bmatrix} u_a^T \\ u_b^T \\ u_c^T \\ u_d^T \end{bmatrix} = \begin{bmatrix} 0 & -1 & 1 \\ .9 & 0 & -1.1 \\ -1 & .5 & .1 \\ .3 & .1 & .1 \end{bmatrix}$$

In figure (2.5) the number of open sets determined by $\{\mathcal{N}_{ab} : a, b \in A\}$ that correspond to strict preference holding between every pair of alternatives in Δ is 18. Examples of preferences that satisfy weak diversity exist where the

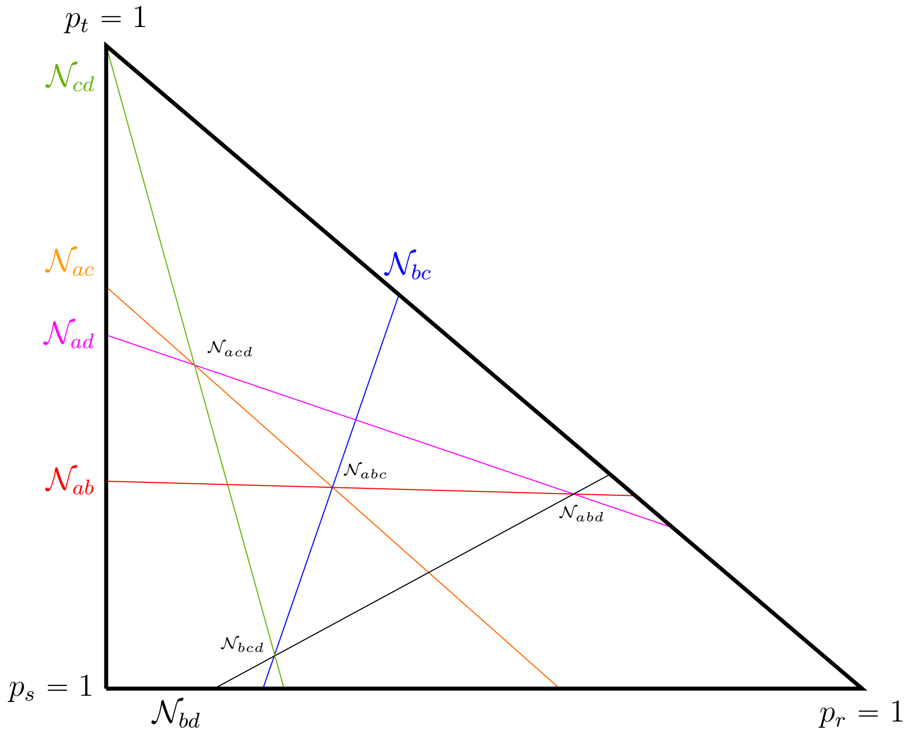


Figure 2.5: The sets $\mathcal{N}_{..}$ and $\mathcal{N}_{...}$ for $|A| = 4$, $|S| = 3$

number of chambers is 16.[6] However when the dimension of the context space is two, 18 is maximal in the sense that any more chambers, given that the sets $\mathcal{N}_{..}$ are flat, signifies a failure of condition (NT). By contrast, when $|A| = 4$, (Div.GS) implies that the number of chambers is at least $4! = 24$.

Similar figures exists for the case where $|A| = 5$. The diagrams also suggest that it is the combination of thinness and flatness of $\{\mathcal{N}_{ab} : a, b \in A\}$ together with (NT) that will surely fail to be observed in experiments: any perturbation of a single set $\mathcal{N}_{..}$ immediately reveals sets where (NT) fails. The next chapter looks at the implications of relaxing the thinness and flatness conditions. As we shall see, diversity is no longer needed to obtain a utility

[6] Azrieli (2011) presents such a figure in a social choice setting.

representation.

Chapter 3

A topological approach to representing context preferences

3.1 Motivation

In this chapter we consider the implications of relaxing the condition of strong betweenness across contexts (p-SB) of chapter 2. Recall that this condition was itself a weakening of the combination axiom of Gilboa and Schmeidler (2003). We do so for two main reasons. The first is that it is a strong condition in its own right, and to the extent that it is the condition that gives rise to an integral or sum in the form of an expected utility or preference representation, it is subject to similar criticisms that other models of expected utility face. For instance, in many situations, as p varies across Δ, preferences $(A, \succ_p)$ may well vary in such a way that is non-linear, possibly exhibiting complementarity across states. Before presenting such an example we look at some more elementary weakening of the condition

Example 3.1 (Getting to university continued)**.** *Recall that in our "get-*

ting to university" example, the state-space S is the union of the three states r:="rain", s:="sunshine" and t:="icy/treacherous", and that Val presently seeks to articulate preferences contingent upon each probability/context that may arise when the resolution of uncertainty become imminent. Suppose that Val is willing to allow for some decisions to be made on the day. That is, in formulating her plan, she is not so obsessive in her plan as to say, for instance: that whenever $p_t > \frac{1}{2}$, she would always choose b:="bus" over c:="cycle" and vice versa whenever $p_t < \frac{1}{2}$, preferences that would admit an expected utility function with threshold set

$$\mathcal{N}_{ab} := \{p \in \Delta(r, s, t) : p_t = p_r + p_s\}.$$

To see this, note that since $p_r + p_s + p_t = 1$, if we solve these two equations by substituting in for, say, p_r, we obtain the solution set

$$\{(p_s, p_t) \in \mathcal{R}^2 : p_t = \frac{1}{2},\ 0 \leqslant p_s \leqslant 1\}.$$

Instead, suppose Val is willing to allow for a thick threshold set $\mathcal{N}_{bc}$ in Δ that separates $\mathcal{B}_{bc}$ from $\mathcal{W}_{bc}$. Suppose, that at the time of making the plan, she only knows that she will choose b over c whenever $p_t > \frac{2}{3}$ and c over b whenever $p_t < \frac{1}{3}$. For p_t between these values she is willing to concede that she is indecisive, and would rather decide on the day.

In this example, since $\mathcal{N}_{ab}$ has open interior in Δ, it cannot be the intersection of a hyperplane with Δ. Also, since $\mathcal{B}_{bc}$ and $\mathcal{W}_{bc}$ are both nonempty, proposition (2.1) of chapter 2 implies that Val's preferences can never have an expected utility or preference representation. However, it is apparent that a pair of expected preference representations, one for each of the sets

$$\{p \in \Delta : p_t = \frac{1}{3}\} \quad \text{and} \quad \{p \in \Delta : p_t = \frac{2}{3}\},$$

would do.

On other hand, if the alternative set is of cardinality greater than two, and Val is willing to accept that indecisiveness is transitive in itself and is dominated by strict preference (negative transitivity), then a utility function representation still ought to be possible. This is the question we pursue in this chapter.

It is plausible that nonlinear models of context-free preferences such as the multiple prior models of Bewley (1987) and Gilboa and Schmeidler (1989) may, when transposed to the context-dependent setting, be able to capture the behaviour present in the above example. We highlight this as a direction for future research.

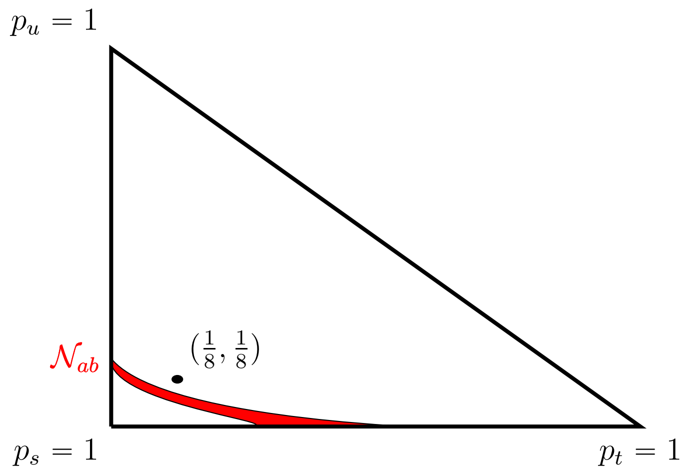
Note that this example also raises the question of how we can represent context preferences that fail to satisfy negative transitivity at each context, but instead accommodate weakened versions that allow for intransitive indecisiveness. Examples of such representations for context-free preferences in the vNM paradigm are the semi-order and interval order models of Fishburn (1968) and Nakamura (1988) respectively. Whilst this was one of the original goals of the present research, we have found that much work still had to be done on the basic model where preferences give rise to a utility function over contexts. We believe the model and techniques we present in this chapter provide some of the preliminary but necessary steps towards more realistic and general models.

Example 3.2. *Building on our flood example of chapter 1, suppose that in addition to the future threat of a flood, there is also the future threat of a mud-slide. For the sake of simplicity, suppose that mud-slides only happen when there is a flood. That is the state-space S is the union of the three events $\{s\}$:="no flood", F:="flood" and $\{u\}$:="mudslide"; where $\{u\} \subset F$. We can write S in terms of the three states s, t and u, where $\{t\} := F \backslash \{u\}$ is the event "flood, but no mud-slide".*

Recall that our decision-maker presently seeks to articulate preferences contingent upon each probability/context that may arise when the threats become imminent. Suppose that one course of action, such as "evacuate to site b" is preferable to "stay at site a" when a flood is certain, and when there is no threat a is preferable to b.

Let the data that we have on preferences be of the following incomplete form. When there is a very low chance of a mudslide, that is $p_u \approx 0$, where $\approx$ means "approximately", the decision-maker is indecisive when $\frac{1}{4} \lessapprox p_t \lessapprox \frac{1}{2}$. Whereas when $p_t \approx 0$, she is indecisive only when $p_u \approx \frac{1}{6}$. Now if $\mathcal{N}_{ab}$ is convex, then a simple calculation shows that the set $\{(p_t, p_u) \in \Delta : p_u = \frac{1}{6} - \frac{2}{6}p_t\}$ should lie in $\mathcal{N}_{ab}$. That is, when there is approximately a one quarter chance of a flood with an equal conditional probability of mudslide and no mudslide ($p_u = \frac{1}{8} = p_t$), the planner should be indecisive. However, it seems plausible that the planner is decisive in favour of evacuating in this context given the unpredictability of a mudslide in addition to the complimentary threat of a flood without a mudslide.

Such preferences might well be of the following form:



In this example, perhaps because floods progress more predictably and have more predictable consequences than mudslides, the decision-maker who is presently putting together an evacuation plan in advance of the threat is willing to allow for more indecisiveness when mudslides are ruled out. However as a mudslide becomes more likely, the unpredictable nature of mudslides, not only in terms of the consequences, but also the abruptness of its arrival forces the planner to be decisive. There is also a degree of complementarity between the two states, an effect which is reflected in the strict convexity of the set of contexts such that b dominates a.

Clearly, in this example, the plan was put together with some broad rules of thumb and there would be many possible preference contours in context space that the above data will permit. However, what is clear is that expected utility of the kind explored in chapter 2 will not do, for $\mathcal{N}_{ab}$ is neither thin, nor convex, nor in fact is $\mathcal{B}_{ab}$. Instead, what is called for is a basic model of context preferences that parallels the continuous ordinal utility representation for context-free preferences of Debreu (1954) and (1964). Once this has been obtained, we may impose additional conditions that impose some structure on how preferences vary with context and which capture features of the preferences described in the above examples.

The structure of the present chapter will accord with these objectives. First, in section (3.2) we present a continuous, ordinal utility representation of complete and transitive preferences that are continuous across the context space. Perhaps more importantly, we identify the properties that a context space needs to have if such a representation is to hold for all possible continuous preferences. Second, in section (3.3), we restrict the class of preferences we seek to model on the grounds that the remaining class is more typical of what we observe in choice under uncertainty. In this section we show that once convexity (in the form of condition (p-SB)) is weakened, we no longer need to consider condition (Div.) and instead may focus on obtaining a representation of preferences that captures behaviour that seems plausible in a wide variety of decision problems.

3.2 A continuous representation of preferences for perfectly normal context spaces

Definition (Asymmetry (Asy.)).

For all $a, b \in A$, $x \in X$: *if* $a \succ_x b$ *then* $\neg(b \succ_x a)$; *equivalently,*

$$x \in \mathcal{B}_{ab} \Rightarrow x \notin \mathcal{W}_{ab}.$$

Definition (Continuity (C'ty)).

For all $a, b \in A$, $x \in X$: if $a \succ_x b$, then there exists an open neighbourhood O of x in X such that for every $x \in O$ we have $a \succ_q b$.

As we discuss in the previous chapter, unlike models such as that of vNM or others that do not define context preferences, here continuity has intuitive appeal because it represents a stability property of strict preferences. That is stability with respect to perturbations of the context as characterized by $p \in \Delta$.

Definition (Negative transitivity (NT))**.**
For all $a, b, c \in A$ and $x \in X$: if $\neg(a \succ_x b)$ and $\neg(b \succ_x c)$, then $\neg(a \succ_x c)$; equivalently, the contrapositive is

$$x \in \mathcal{B}_{ac} \quad \Rightarrow \quad x \in \mathcal{B}_{ab} \cup \mathcal{B}_{bc}$$

It is well known that (NT) and (Asy) of the relation $\succ$ are together equivalent to assuming that the union of $\succ$ and $\sim$, $\gtrsim$ is both complete and transitive–see Fishburn (1979) ch.2 for instance. In this chapter we will work with strict preferences $\succ$, however where necessary we will make use of the weak preference relation $\gtrsim$.

We now state and prove a very general existence theorem for context preferences. It is in some sense the most general theorem we can hope for if we seek a representation that varies continuously across contexts. This is because of the following theorem on spaces that are *perfectly normal* ie. normal spaces X for which every closed set can be written as a countable intersection of sets that are open in X (Munkres p.229). We recall that the space X is said to be *normal* if for every pair of disjoint closed subsets of X there exist disjoint open sets containing A and B respectively.

Theorem 3.1 (Michael's selection theorem). *A space X is perfectly normal if and only if whenever $g, h : X \to \mathcal{R}$ are upper (resp. lower) semi-continuous and $g \leqslant h$, there is a continuous $f : X \to \mathcal{R}$ such that $g \leqslant f \leqslant h$ and $g(x) < f(x) < h(x)$ whenever $g(x) < h(x)$.*

This equivalence is relevant for preferences that are indexed by elements of a context space because if the context space is not perfectly normal then *there exist* $g, h : X \to \mathcal{R}$ that are upper (resp. lower) semi-continuous and $g \leqslant h$, such that *for no* continuous $f : X \to \mathcal{R}$ do we have $g \leqslant f \leqslant h$ and $g(x) < f(x) < h(x)$ whenever $g(x) < h(x)$. This as we will see in the proof of our theorem and subsequent discussion, would imply that there exists context preferences $\{(A, \succ_x) : x \in X\}$ such that there is no representation $U : A \times X \to \mathcal{R}$.

We will not provide a comprehensive answer to the question of how restrictive the condition that the context space be perfectly normal is at this point. Though we note for countable S the simplex of probability measures Δ on S is a subspace of $\mathcal{R}^S$, which is a Hilbert space under the standard Euclidean metric, so that by Steen and Seebach p.65, Δ is a metric space. Then by Munkres p. 229 every metrizable space is perfectly normal. Thus for countable S, Δ is a suitable context space. This is not true for uncountable S, by counterexample 105, p125 of Steen and Seebach.

Another example of a context space in the literature is any countable product of the discrete space of non-negative integers with the Cartesian product topology as is found in the case-based decision theory of Gilboa and Schmeidler. By Steen and Seebach p.121, this is a complete metric space, and so this too is a suitable space of contexts. On the other hand, uncountable products of the nonnegative integers with the product topology are, by counterexample 103 of Steen and Seebach, not normal spaces, and so they may be unsuitable depending on preferences.

The representation we are seeking is of the following form.

Definition 3.1. $U : A \times X \to \mathcal{R}$ *is said to be a* continuous context utility representation *of preferences* $\{(A, \succ_x) : x \in X\}$ *if it is continuous and for all* $a, b \in A$ *and* $x \in X$*,*

$$a \succ_x b \quad \Leftrightarrow \quad U(a, x) > U(b, x).$$

A continuous (context) representation U is said to be ordinal *if for any other representation V of preferences, there exists a family of strictly increasing functions $\{f_x : \mathcal{R} \to \mathcal{R}\}_{x \in X}$ such that for each x, $V(\cdot, x) = f_x \circ U(\cdot, x)$.*

Remark 3.1. *Note that in definition (3.1), the continuity of $f_.$ across contexts is implied by the continuity of U and V.*

We now state and prove our representation theorem for context preferences that satisfy the aforementioned conditions.

Theorem 3.2. *Let A be countable and X be a perfectly normal space of contexts. The following two statements are equivalent.*

1) *Context preferences $\{(A, \succ_x) : x \in X\}$ satisfy (Asy.), (C'ty) and (NT).*

2) Context preferences have a continuous ordinal utility representation.

Proof. Let $\{1, 2, 3 \ldots\}$ be an arbitrary enumeration of A, and by $[j]$ we will denote the subset of A that contains the first j elements of the enumeration. By $U^j : [j] \times X \to \mathcal{R}$ we will denote the utility representation of the projection of preferences $\{(A, \succ_x) : x \in X\}$ onto the first j elements of the enumeration. That is, if we recall that for each $x \in X$, $\succ_x$ is a subset of $A \times A$, then we see that $\{\succ_x : x \in X\} \subset (A \times A)^X$. Hence by the projection of preferences onto $[j]$ we mean

$$\{\succ_x : x \in X\} \cap \big([j] \times [j]\big)^X$$

which is a well defined intersection since

$$\begin{aligned}
\big([j] \times [j]\big)^X \cap (A \times A)^X &= \prod_{x \in X} \big([j] \times [j]\big) \cap (A \times A) \\
&= \prod_{x \in X} ([j] \cap A) \times ([j] \cap A) \\
&= \big([j] \times [j]\big)^X.
\end{aligned}$$

We will use this projection to proceed by induction on A. Thus we first show that a continuous representation for the basic case: which we take to be $j = 1$.

Let $U^1(1, x) \equiv 0$ for all $x \in X$. By condition (Asy.), U^1 is a representation for the projection of preferences onto $[1]$ and it is clearly continuous. This completes the proof for the basic case. Now suppose that for some $j \in \mathbb{N}$ greater than 1 there exists a representation U^{j-1} of the projection onto $[j-1]$. We need to show that the conditions imply the existence of a representation of the projection onto $[j]$.

For $a \in [j-1]$ we set $U^j(a,\cdot) = U^{j-1}(a,\cdot)$. Then by the induction hypothesis, for all $a, b \in [j-1]$ and $x \in X$ we have,

$$a \succ_x b \quad \Leftrightarrow \quad U^j(a,x) > U^j(b,x),$$

and on this subset of $[j]$, U^j is continuous.

We now need to select a continuous function $U^j(j,\cdot)$ on X such that for each x, $U^j(\cdot,x) : [j] \to \mathcal{R}$ represents $\succ_x \cap [j] \times [j]$. We will do this by first defining the lower and upper envelopes, $\underline{U}$ and $\overline{U}$ respectively of $U^j([j-1],X)$ relative to alternative j. That is, informally speaking, for the lower envelope relative to j we seek the function whose graph is the set of points $\{(x, U^j(\cdot,x)) : x \in X\}$ that can be seen by looking down from the position of j in the preference order at context x. (Similarly, the upper envelope relative to j, it is the set of points that can be seen by looking up.)

We then show that these two functions are respectively upper and lower semi-continuous and that the latter dominates the former pointwise. This, together with the fact that X is perfectly normal implies, via Michael's selection theorem that the required function $U^j(j,\cdot)$ exists.

Recalling that $\mathcal{W}_{j[j-1]} := \bigcap_{k=1}^{j-1} \mathcal{W}_{jk}$ is the set of elements of X such that j is strictly worse than all the elements of $[j-1]$, and similarly for $\mathcal{B}$, let us now define the lower and upper envelopes of U^j relative to j.

$$\underline{U}(x) := \begin{cases} \min_{k\in[j-1]} \left\{U^j(k,x)\right\} - 1 & \text{if } x \in \mathcal{W}_{j[j-1]} \\ \max_{k\in[j-1]} \left\{U^j(k,x) : j \gtrsim_x k\right\} & \text{otherwise.} \end{cases}$$

This function is well defined for the following reasons: firstly, $[j-1]$ is compact; secondly, for each x in

$$X \backslash \mathcal{W}_{j[j-1]} = \bigcup_{k=1}^{j-1} \left(\mathcal{B}_{jk} \cup N_{jk}\right)$$

there exists $k \in [j-1]$ such that $j \gtrsim_x k$; and thirdly, the fact that $\gtrsim_x$ is complete and transitive means that if D_x is the subset of alternatives in $[j-1]$ that j weakly dominates at x, there is at least one element of D_x that is maximal in D_x under $\gtrsim_x$. Similarly, the upper envelope of $U^j([j-1], X)$ relative to j is also well defined as follows:

$$\overline{U}(x) := \begin{cases} \max_{k\in[j-1]} \left\{U^j(k,x)\right\} + 1 & \text{if } x \in \mathcal{B}_{j[j-1]} \\ \min_{k\in[j-1]} \left\{U^j(k,x) : k \gtrsim_x j\right\} & \text{otherwise.} \end{cases}$$

Claim 3.1. *For all $x \in X$, $\underline{U}(x) \leqslant \overline{U}(x)$.*

Proof. With a view to obtaining a contradiction, suppose that for some $x \in X$, $\overline{U}(x) < \underline{U}(x)$. Then by definition x cannot be in the union of $\mathcal{W}_{j[j-1]}$ and $\mathcal{B}_{j[j-1]}$. Thus for some $k, l \in [j-1]$, $\underline{U}(x) = U^j(k,x)$ and $\overline{U}(x) = U(l,x)$. Now once more by definition,

$$l \gtrsim_x j \gtrsim_x k,$$

so that by condition (NT) it follows that: $l \gtrsim_x k$. Then since U^j is equal to U^{j-1} on $[j-1] \times X$, which, by the induction hypothesis, represents the projection of preferences onto $[j-1]$, we have $U(l,x) \geqslant U(k,x)$. This is the required contradiction, and so we see that $\underline{U}$ is pointwise weakly dominated by $\overline{U}$. $\square$

Purely for expositional purposes, we introduce two fictional alternatives $\underline{a}$ and $\overline{a}$, such that for all $x \in X$ and $k \in [j]$, we have $\overline{a} \succ_x k \succ_x \underline{a}$. Accordingly, we define $[j-1]^* := [j-1] \cup \{\underline{a}, \overline{a}\}$, and for each $x \in X$, let

$$U^j(\underline{a}, x) \equiv \min_{k \in [j-1]} \{U^j(k,x)\} - 1$$

and

$$U^j(\overline{a}, x) \equiv \max_{k \in [j-1]} \{U^j(k,x)\} + 1.$$

Now define the lower envelope of $U^j([j-1]^*, X)$ relative to j, and note that for all $x \in X$ there exists $k \in [j-1]^*$ such that $j \succsim_x k$, so that

$$\underline{U}(x) := \max_{k \in [j-1]^*} \left\{U^j(k,x) : j \succsim_x k\right\}$$

is well defined and similarly, so is

$$\overline{U}(x) := \min_{k \in [j-1]^*} \left\{U^j(k,x) : k \succsim_x j\right\}.$$

Now, by construction, the lower and upper envelopes of $U^j([j-1], X)$ relative to j are respectively identical to the lower and upper envelopes of $U^j([j-1]^*, X)$ relative to j. Moreover, we can readily see that for all $x \in X$ if for some $k \in [j-1]^*$, $k \sim_x j$, then $\underline{U}(x) = U^j(k,x) = \overline{U}(x)$; on the otherhand, if $\underline{U}(x) = \overline{U}(x)$ for some x, then there exists $k \in [j-1]$, for it cannot be $\underline{a}$ or $\overline{a}$, such that $k \sim_x j$. So we see that there exists $\underline{U}(x) = \overline{U}(x)$ if and only if there exists $k \in [j-1]$ such that $k \sim_x j$. Equivalently, for each $x \in X$, if $\underline{U}(x) = U^j(g, x)$ and $\overline{U}(x) = U^j(l, x)$ for some $g, l \in [j-1]^*$, then

$$\overline{U}(x) > \underline{U}(x) \quad \Leftrightarrow \quad l \succ_x g,$$

and for all other $k \in [j-1]^*$, either $k \succ_x l$ or $g \succ_x k$. This, together with claim (3.1), shows that provided $\underline{U}$ is usc and $\overline{U}$ is lsc, Michael's selection theorem tells us that the desired continuous function exists.

We now show that $\underline{U}$ is upper semi continuous (usc). We do so by showing that $\underline{U}$ is infact continuous everywhere except the contexts where alternative j changes from being strictly worse to being indifferent to some other alternative(s) and the set of alternatives that j dominates is unchanged. At such points, $\underline{U}$, should, intuitively speaking, increase because the set of elements of $[j-1]^*$ that are weakly dominated by j at x will have increased in cardinality by at least one, and we recall that $\underline{U}$ is defined in terms of the maximum over D_x. The fact that the alternative space is discrete and strict preference is continuous then implies that there is a jump up in the value of $\underline{U}$ in such contexts, a property that is satisfied by usc functions.

We now provide a formal proof of this argument.

Claim 3.2. *On X, $\underline{U}$ is upper semicontinuous and $\overline{U}$ is lower semicontinuous.*

Proof. For each $x \in X$, let D_x^* be the set of elements of $[j-1]^*$ that are weakly dominated by j at x. Similarly, let $E_x^* := \{k \in [j-1]^* : k \succsim_x j\}$. Condition (Asy.) and the definition of $\sim_.$ imply that for all $x \in X$ and $k \in [j-1]$, exactly one of following relationships must hold:

$$j \succ_x k, \quad k \succ_x j \quad \text{or} \quad k \sim_x j.$$

So that for each x, we have $D_x^* \cup E_x^* = [j-1]^*$. Furthermore, $l \in D_x^*$, $k \in E_x^*$ implies that $l \gtrsim_x k$ by transitivity of $\gtrsim_x$, and $l \sim_x k$ holds if and only if $D_x^* \cap E_x^* \neq \emptyset$, which in turn is true, if and only if $x \in \mathcal{N}_{jk}$ for some $k \in [j-1]$. When $D_x^* \cap E_x^*$ is empty, therefore, the two sets form a partition of $[j-1]^*$.

We now prove the claim that $\underline{U}$ is upper semicontinuous.

Let $x \in X$ be such that for any open set O that contains x and is sufficiently small, we have $D_z^* = D_x^*$ for all $z \in O$. Thus for each $z \in O$, $\underline{U}(z) = U^j(k, z)$ for some $k \in D_x^*$ and so

$$O = \bigcup_{k \in D_x^*} \bigcap_{l \in D_x^*} \{x \in O : k \gtrsim l\} = O \cap \big(\bigcup_{k \in D_x^*} \Delta \backslash \mathcal{W}_{kD_x^*} \big).$$

By condition (C'ty) and the fact that the finite intersection of open sets is open, we see that $\mathcal{W}_{kD_x^*}$ is open in X and so its complement is closed. Therefore in the subspace topology of O, $O \cap X \backslash \mathcal{W}_{kD_x^*}$ is also closed. Now on each of the sets $O \cap \Delta \backslash \mathcal{W}_{kD_x^*}$, $\underline{U} = U^j(k, \cdot)$, and by the induction hypothesis, $U^j(k, \cdot)$ is continuous on X and hence continuous on each of its subsets, so in particular the restriction of $U^j(k, \cdot)$ to $O \cap \Delta \backslash \mathcal{W}_{kD_x^*}$ is a continuous function.

In the intersection of any pair of sets $\Delta \backslash \mathcal{W}_{kD_x^*}$, $\Delta \backslash \mathcal{W}_{lD_x^*}$, $k, l \in D_x^*$ it is clear that by condition (Asy.) we cannot have strict preference in either direction between k and l. Thus for any $z \in O$ in such an intersection we have:

$$U^j(k,\cdot) = \underline{U}(\cdot) = U^j(l,\cdot).$$

In the same way, for any $z \in O$ and $D \subset D_x^*$ such $\bigcap_{k \in D} \Delta \backslash \mathcal{W}_{kD_x^*}$ we have $U^j(k,z) = \underline{U}(z)$ for all $k \in D$.

Now let C be any closed subset of $\mathcal{R}$. Then we have

$$O \cap \underline{U}^{-1}(C) = O \cap \big(\bigcup_{k \in D_x} U^j(k,\cdot)^{-1}(C) \big).$$

Continuity of $U^j(k,\cdot)$ for each $k \in D_x$ implies that $U^j(k,\cdot)^{-1}(C)$ are closed in X, and the fact that D_x is finite implies that the union on the right-hand-side of this equation is closed in X. Hence, in the subspace topology, both sides are closed in O and therefore $\underline{U}$ is continuous at x.[1]

By condition (C'ty), the above argument accounts for $x \in \mathcal{B}_{kD} \cap \mathcal{W}_{kE}$, for each partition D, E of $[j-1]^*$. That is, all points x such that $j \not\sim_x k$ for any $k \in [j-1]$. On the interior of the set $\bigcup_{k=1}^{j-1} \mathcal{N}_{jk}$ relative to X, there exists $k \in [j-1]$ and an open neighborhood O of x that is contained in $\mathrm{Int}_X \mathcal{N}_{jk}$. As such, on O, $\underline{U}$ is equal to the continuous function $U^j(k,\cdot)$ and is thereby continuous.

It remains for us to consider $x \in \mathrm{bd}_X \mathcal{N}_{jk}$ for arbitrary $k \in [j-1]$. Here, in every open neighborhood O of x, there exists $y \in O$ such that $y \in \mathcal{B}_{jk} \cup \mathcal{W}_{jk}$.

[1]This argument is based on what is called "The pasting lemma" (Munkres p.124).

Let $\{O_n : n \in \mathbb{N}\}$ be a sequence of open sets that contain x which satisfies $\bigcap_{n=1}^{\infty} O_n = \{x\}$. (Such a sequence exists precisely because X is perfectly normal.)

We will consider sequences of contexts $\{y_n\}$ such that $y_n \in O_n$ for each n, so that $\lim_n y_n = x$. We do so by first partitioning each set O_n into the three sets

$$O_n \cap \mathcal{N}_{jk}, \quad O_n \cap \mathcal{W}_{jk} \quad \text{and} \quad O_n \cap \mathcal{B}_{jk}.$$

Any such sequence of contexts $\{y_n\}$ that has infinitely many elements that lie in more than one of the three sets that determine the partition, say $\mathcal{B}_{jk}$ and $\mathcal{W}_{jk}$, has convergent subsequences $\{y'_{n_m}\}$ and $\{y''_{n_m}\}$ that lie wholly in $\mathcal{B}_{jk}$ and $\mathcal{W}_{jk}$ respectively. So it suffices to consider sequences in each of the partitions separately.

For each sequence $\{y_n \in O_n \cap \mathcal{N}_{jk}\}$ we have $\underline{U}(y_n) = U^j(k, y_n)$ for each $n \in \mathbb{N}$, and by the induction hypothesis $U^j(k, \cdot)$ is continuous, so $\underline{U}(y_n)$ converges to $\underline{U}(x)$.

For the sequence of sets $\{O_n \cap \mathcal{W}_{jk}\}$, we note that $z \in O_n \cap \mathcal{W}_{jk}$ implies that $k \succ_z j$, and, by condition (NT), for all $l \in D_z$ we have $k \succ_z l$, so that $U^j(k, z) > \underline{U}(z)$. Let us consider the following two exhaustive cases.

Case 1. There exists $n \in \mathbb{N}$ such that for all $z \in P_n \equiv \bigcup_{m \geqslant n}(O_m \cap \mathcal{W}_{jk}) \cup \{x\}$: if $l \in D_z$, then $k \succ_x l$. That is, there exists an open neighborhood Q of x such that for all $y \in (Q \cap \mathcal{W}_{jk}) \cup \{x\}$, $U^j(k, y) > \underline{U}(y)$. Thus, for any sequence of contexts $\{y_m \in O_m \cap \mathcal{W}_{jk} : m \in \mathbb{N}\}$, $\{y_m\}$ converges to $x \in \mathrm{bd}_X \mathcal{W}_{jk}$ and moreover,

$$\limsup_m \underline{U}(y_m) < U(x) = U^j(k, x)$$

Thus in this case, $\underline{U}$ is x is upper semicontinuous.

Case 2. (This is the negation of Case 1.) That is for all open neighborhoods Q of x, there exists $z \in (Q \cap \mathcal{W}_{jk}) \cup \{x\}$ with $U^j(k, z) \leqslant \underline{U}(z)$. Now if $z \neq x$, then $z \in \mathcal{W}_{jk}$ and $U^j(k, z) \leqslant \underline{U}(z)$ imply $j \gtrsim_z l \gtrsim_z k \succ_z j$, which contradicts condition (NT). Thus $z = x$ is the only context such that for all $n \in \mathbb{N}$ there exists $z \in P_n$ and $l \in D_z$ such that $l \gtrsim_x k$. Once more by condition (NT) we see that for such l, $l \sim_x k$.

Now note that for each $y_n \in O_n \cap \mathcal{W}_{jk}$ we have $U^j(l, y_n) \leqslant \underline{U}(y_n) < U^j(k, y_n)$; this is equivalent to

$$0 \leqslant \underline{U}(y_n) - U^j(l, y_n) < U^j(k, y_n) - U^j(l, y_n).$$

Then using the triangle inequality on $\mathcal{R}$, the above bound, and the fact that $U^j(k, x) = \underline{U}(x) = U^j(l, x)$ we have:

$$\begin{aligned}
0 &\leqslant |\underline{U}(y_n) - \underline{U}(x)| \\
&\leqslant |\underline{U}(y_n) - U^j(l, y_n)| + |U^j(l, y_n) - \underline{U}(x)| \\
&\leqslant |U^j(k, y_n) - U^j(l, y_n)| + |U^j(l, y_n) - \underline{U}(x)| \\
&\leqslant |U^j(k, y_n) - \underline{U}(x)| + 2|U^j(l, y_n) - \underline{U}(x)|.
\end{aligned}$$

Thus $|\underline{U}(y_n) - \underline{U}(x)|$ converges to 0 by continuity of $U^j(k,\cdot)$ and $U^j(l,\cdot)$ via the induction hypothesis. This completes the proof of Case (2). Indeed, in this case, if for instance $\mathcal{B}_{jk}$ is empty, $\underline{U}$ is in fact continuous at the boundary of $\mathcal{W}_{jk}$.

We now turn to the remaining sequence of sets $\{O_n \cap \mathcal{B}_{jk}\}_{n\in\mathbb{N}}$. Now as above, for each n, let $P_n \equiv \bigcup_{m\geqslant n}(O_m \cap \mathcal{B}_{jk}) \cup \{x\}$. Now define the sequence of subsets of $[j-1]^*$

$$F_n := \bigcup_{y\in P_n} D_y,$$

then because $F_n = F_{n+1} \cup \bigcup_{y\in O_n\cap\mathcal{B}_{jk}} D_y$, F_n is a decreasing sequence of sets: each of which contains the alternative k. Now since $[j-1]^*$ is a finite set, there exists $n' \in \mathbb{N}$ such that for all $n \geqslant n'$, $F_n = F_{n'} \equiv F$.

Once more there are two cases. The first holds when there exists $n' \in \mathbb{N}$ such that $y, z \in P_{n'}$ implies $D_y = D_z$. By the construction of $\{P_n\}$, this implies that the same is true for all $n \geqslant n'$. In this case we may use the partition lemma argument on $P_{n'}$ to show that the restriction of $\underline{U}$ to P'_n is continuous. Although, $P_{n'}$ is not a neighborhood of x, it is clear that for any sequence $\{y_n\} \subset P_{n'}$ that is converging to x, we have $\underline{U}(y_n)$ converges to $\underline{U}(x)$.

The remaining case is where for all $n \in \mathbb{N}$, there exists $y, z \in P_n$ such that $D_y \neq D_z$. Let us define the set

$$L := \bigcap_n \{l \in F : l \in D_y \backslash D_z, \text{for some } y, z \in P_n\}.$$

The fact that $P_n \subset \mathcal{B}_{jk} \cup \{x\}$ implies that $k \in D_y \cap D_z$ for all $y, z \in P_n$, thus $F \neq L_n$ for all n.

Suppose that $l \in L$, and $l \succ_x j$. Then by condition (C'ty) there exists an open neighborhood Q of x such that $l \notin D_y$ for all $y \in Q$. Now for n sufficiently large, the fact that $\bigcap_n O_n = \{x\}$ implies that $P_n \subset Q$. Thus, either $P_n \cap Q = \{x\}$ (if we are considering the discrete topology on X), or, by definition of L, there exist $y, z \in Q$ such that $l \in D_y \backslash D_z$. In either case we obtain a contradiction: in the former it is the fact that we have found $n \geqslant n'$ such that $D_y = D_z$ for all $y, z \in P_n$; and in the second, it is the definition of Q.

Now suppose that $j \succ_x l$. In this case there exists an open neighborhood Q of x such that $l \in D_y$ for all $y \in Q$. In this case, we see that like k, $l \in F \backslash L$.

Thus, by default, we see that $l \sim_x j$. Now let $\{z_n\}$, $z_n \in P_n$ be such that $l \in D_{y_n} \backslash D_{z_n}$ for some $y_n \in P_n$. Thus, for all z_n we have $l \succ_{z_n} j$. Now since z_n converges to x, we know that for every open neighborhood Q of x there exists n such that $z_n \in Q$. Thus, $x \in \mathrm{bd}_X \mathcal{W}_{jl}$. This case is therefore identical to Case (2) above with the roles of k and l reversed.

This completes the proof of the fact that $\underline{U}$ is upper semicontinuous on X. The proof that $\overline{U}$ is lower semicontinuous is identical except for the fact that the points of discontinuity lie on the boundary of $\mathcal{B}_{j\cdot}$ rather than $\mathrm{bd}_X \mathcal{W}_{j\cdot}$ as we have shown to be the case for $\underline{U}$. $\square$

The proof of this last claim completes the proof of the theorem. $\square$

3.3 The topological approach to uncertainty

The main strength of the representation of section (3.2) is that it is general. That is, it allows for a very diverse class of context preferences. It clearly encompasses the expected utility representations of chapter 2, but also the transitive expected preference representations of that chapter. However for some decision problems it may be too general to be of any use in applications. That is, whilst continuity imposes some useful and realistic structure across contexts, we may be able to find additional conditions on preferences that are reasonable in certain decision problems.

As an example, when the context space is the set of probability distributions Δ on $S = \{s, t, u\}$, and preferences are such that for some $a, b \in A$ and $p = (p_s,\, p_t,\, p_u) \in \Delta$, we have $a \succ_p b$ and $a \succ_s b$ where $p \neq \delta_s$. (We recall that δ^v is the basis vector in $\mathcal{R}^{|S|}$ corresponding to the axis that we have assigned to each state $v \in S$, and $\succ_s$ is short hand for $\succ_{\delta_s}$.)

Now consider how preferences change as the probability of state s increases and the relative weights of each of the other states remains the same. In other words, suppose that for $\lambda \in [0, 1]$ we look at preferences at contexts r in $\mathrm{conv}(p, \delta_s)$, that is

$$\begin{aligned} r &= \lambda\delta^s + (1-\lambda)p \\ &= p_s\,\delta^s + (1-\lambda)p_t\,\delta^t + (1-\lambda)p_u\,\delta^u. \end{aligned}$$

Now it may seem unreasonable that there exist arbitrarily many such probability distributions r such that $b \gtrsim_r a$. However the representation in section (3.2) is general enough to accommodate such preferences. Indeed if the objective is to elicit preferences, the generality of that model might well prove too costly in terms of number of decisions needed to do so.

Indeed we may often be willing to go further and commit to the following condition.

Definition (Betweenness with respect to states (s-B))**.** *Both*

> (s-B$\succ$): *for all* $a, b \in A$, $p \in \Delta$ *and* $s \in S$*: if* $a \succ_p b$ *and* $a \succ_s b$*, then for* $q = \lambda p + (1-\lambda)\delta_s$, $0 < \lambda < 1$ *we have* $a \succ_q b$; *and,*
>
> (s-WB): *condition* (s-B$\succ$) *with the relation* $\gtrsim_.$ *replacing strict preference.*
>
> *Equivalently, both*
>
> $$p, \delta_s \in \mathcal{B}_{ab} \Rightarrow \operatorname{conv}(p, \delta_s) \subset \mathcal{B}_{ab}\,;\ \textit{and}$$
> $$p, \delta_s \in \mathcal{B}_{ab} \cup \mathcal{N}_{ab} \Rightarrow \operatorname{conv}(p, \delta_s) \subset \mathcal{B}_{ab} \cup \mathcal{N}_{ab}.$$

Condition (s-B$\succ$) is substantial weakening of condition (p-SB) (strong betweenness across contexts) which was introduced in chapter 2. Condition (p-SB) implies (s-B$\succ$) as well as the following condition which is much more difficult to motivate.

Definition (Thinness of no-strict preference (T))**.**

For all $a, b \in A$, $p \in \Delta$ *and* $s \in S$*: if* $a \sim_p b$ *and* $a \succ_s b$*, then* $a \succ_r b$ *for all*

$r = \lambda p + (1-\lambda)\delta_s$ such that $0 < \lambda < 1$; equivalently,

$$\mathrm{Int}_\Delta \mathcal{N}_{ab} = \varnothing.$$

This condition genuinely implies schizophrenic changes in preferences as contexts vary near what, in the presence of this condition, may justifiably be called the threshold set $\mathcal{N}_{ab}$. This condition is closely related to the strong part of the sure-thing principle of Savage (1972). We will only make use of condition (T) at the end of this chapter when we relate the model proposed here to that of chapter 2.

In chapter 2 we also introduced the following condition which appears in one form or another in a majority of the literature on choice under uncertainty: for instance it is the weak part of the "sure thing principle" in Savage (1972) or "monotonicity" in Schmeidler (1989). We refer to it as weak the weak sure-thing principle as this seems to be the most descriptive term available whereas monotonicity is misleading and, like the "independence" condition, it carries different meanings in different fields. The weak sure-thing principle is akin to weak Pareto optimality as used in the literature on social choice: see for instance Blackorby, Donaldson and Weymark (1984) or d'Aspremont and Gevers (2002).

Definition (Weak sure-thing principle (WP))**.**
For all p in Δ : if $\neg(a \succ_s b)$ for all $s \in \mathrm{supp}(p)$, then $\neg(a \succ_p b)$; equivalently,

$$p \in \mathcal{B}_{ab} \;\Rightarrow\; \exists s \in \mathrm{supp}(p) : \delta_s \in \mathcal{B}_{ab}.$$

An interesting example of where this condition fails to hold is provided in section 5 of Shafir, Simonson and Tversky (1993).

If the decision maker finds conditions (s-B$\succ$) and (WP) agreeable, then preferences immediately have the property that for any given pair of alternatives $a, b \in A$, the sets $\mathcal{B}_{ab}$ of contexts where strict preference holds for a over b are connected. Before stating and proving this fact we recall the following notions from topology: separation, connectedness, a path and path-connectedness.

Definition 3.2 (Munkres (2000) p164)**.** *A* separation *of a topological space Z is a pair U, V of disjoint nonempty open sets of Z whose union is Z. The space Z is said to be* connected *if there does not exist a separation of Z.*

Definition 3.3 (Munkres (2000) p171)**.** *Given points p and q of a space Z, a* path *in Z from p to q is a continuous map $f : [a, b] \longrightarrow Z$ of some closed interval in the real line into Z, such that $f(a) = p$ and $f(b) = q$. A space is said to be* path connected *if every pair of points in Z can be joined by a path in Z.*

Proposition 3.1. *If for all $p \in \Delta$, $\succ_p$ is asymmetric and continuous, then conditions (s-B$\succ$) and (WP) together imply that for every $a, b \in A$ the set $\mathcal{B}_{ab}$ is a connected set.*

Proof. Fix an arbitrary pair a, b in A. Note that if $\mathcal{B}_{ab} \neq \emptyset$, then by condition (WP): $p \in \mathcal{B}_{ab}$ implies that there exists $s \in \text{supp}(p)$ such that $\delta_s \in \mathcal{B}_{ab}$. Thus $a \succ_s b$. Then by condition (s-B$\succ$), every $p \in \mathcal{B}_{ab}$ is connected to δ_s via a straight-line path in Δ, $[p, \delta_s]$, in Δ. Finally, for every $q \in \mathcal{B}_{ab}$ the union of $[p, \delta_s]$ with $[q, \delta_s]$ is also a path in Δ, so that $\mathcal{B}_{ab}$ is path connected and hence connected. $\square$

In fact condition (s-B$\succ$) in combination with (WP) implies much more than connectedness of the sets $\{\mathcal{B}_{ab}\}$. Indeed they imply that each of these sets is star-shaped or, equivalently, star-convex with respect to every element of the set

$$\Delta_{a \succ b} := \operatorname{conv}\big(\{\delta_s : \delta_s \in \mathcal{B}_{ab}\}\big) = \{p \in \Delta : \text{for all } s \in \operatorname{supp}(p),\, a \succ_s b\},$$

where this term is defined as follows.

Definition 3.4. *In a vector space a set Z is called* star-shaped *or* star-convex *if for all $z \in Z$ there exists $x \in Z$ such that* $\operatorname{conv}(z, x) \subset Z$*. The set of points with respect to which Z is star-convex is called the* convex kernel *of Z.*

The following theorem provides the definition of star-convexity we need.

Theorem 3.3 (Smith (1968))**.** *In a vector space a set Z is star-shaped if and only if the intersection of all the maximal convex sets of Z is nonempty.*

The implications for strict preference of imposing conditions (s-B$\succ$) and (WP) in isolation of the other conditions are summarized in the following statement.

Proposition 3.2. *Preferences $\{(A, \succ_p) : p \in \Delta\}$ satisfy conditions (s-B$\succ$) and (WP) if and only if for each $a, b \in A$, the set $\mathcal{B}_{ab}$ satisfies*

(i) it is star-shaped and $\Delta_{a \succ b}$ lies in its convex kernel;

(ii) $\mathcal{B}_{ab} \cap \Delta_{b \succsim a}$ is empty.

Remark 3.2. *It is easy to see that because*

$$\begin{aligned}\Delta_{b \succsim a} :&= \operatorname{conv}\big(\{\delta_s : \delta_s \in \Delta \backslash \mathcal{B}_{ab}\}\big)\\ &= \{p \in \Delta : \textit{for all } s \in \operatorname{supp}(p),\ \delta_s \in \Delta \backslash \mathcal{B}_{ab}\},\end{aligned}$$

condition (WP) is in fact equivalent to property (ii) in this proposition. Moreover, weakening (WP) to the condition "$p \in \mathcal{B}_{ab}$ implies that there exists $s \in \Delta$ such that $\delta_s \in \mathcal{B}_{ab}$" would, in conjunction with (s-B≻), be equivalent to $\mathcal{B}_{ab}$ satisfying property (i) for all $a, b \in A$.

As we will see in the proof to this proposition, the same argument that is used to prove (i) shows that condition (s-WB) is equivalent to the property that $\Delta \backslash \mathcal{B}_{ab}$ is star-convex with respect to $\Delta_{b \gtrsim a}$. This in turn implies that condition (s-WB) is sufficient for condition (WP).

An alternative approach may be to define preferences only on the interior of the simplex where all distributions assign positive probability to each and every state. We intend to pursue this approach, which is also adopted in the model of Hammond (1988), elsewhere.

Proof. ($\Rightarrow$ (i)) First note that condition (WP) ensures that $\Delta_{a \succ b}$ is nonempty whenever $\mathcal{B}_{ab}$ is. Second, we note that if $p \in \Delta_{a \succ b}$, then $\operatorname{conv}\left(p, \Delta_{a \succ b}\right)$ is equal to $\Delta_{a \succ b}$ and has extremal points $\{\delta_s : \delta_s \in \mathcal{B}_{ab}\}$, and if not, then p is also an extremal point. Thus we only prove that $\Delta_{a \succ b} \subset \mathcal{B}_{ab}$, as the proof of the latter is identical, but for the additional notation to account for p not being extremal in Δ. We do so by induction.

We appeal to the general property of simplices that for $k \geqslant 1$, the $k-1$-simplices define the boundaries of the k-simplices. The set of extremal points of $\Delta_{a \succ b}$, is the set of 0-subsimplices of Δ, are contained in $\mathcal{B}_{ab}$. We denote the

1-simplices of Δ as follows

$$\{\Delta(s,t) : s,t \in S, s \neq t\}.$$

To prove the basic case we note that for every pair of states s, t, since $\delta_s, \delta_t \in \mathcal{B}_{ab}$, condition (s-B$\succ$) implies that $\mathrm{conv}(\delta_s, \delta_t) \equiv \Delta(s,t)$ is a subset of $\mathcal{B}_{ab}$.

The general case is proven by appealing to the induction hypothesis. That is, if $\Delta_{k-1} \subset \mathcal{B}_{ab}$ for every (k-1)-face of Δ_k, then the boundary of Δ_k is contained in $\mathcal{B}_{ab}$. Moreover, Δ_k is a compact, convex subset of its affine hull, $\mathcal{R}^k$, and as a simplex, it contains an extremal point δ_s. We can therefore apply claim (3.5)–stated and proved in the proof of proposition (3.3) at the end of this section–together with condition (s-B$\succ$) to show that every point in Δ_k is an element of $\mathcal{B}_{ab}$. The above argument implies that for every $p \in \mathcal{B}_{ab}$, conv$(p, \Delta_{a \succ b}$ is a subset of $\mathcal{B}_{ab}$. In turn therefore, $\mathcal{B}_{ab}$ is equal to

$$\bigcup_{p \in \mathcal{B}_{ab}} \mathrm{conv}\left(p, \Delta_{a \succ b}\right),$$

and so every maximal convex set in $\mathcal{B}_{ab}$ is the union $\mathrm{conv}\left(p, \Delta_{a \succ b}\right)$ over $p \in C$ for some $C \subset \mathcal{B}_{ab}$. Then since the set $\Delta_{a \succ b}$ is nonempty and equal to

$$\bigcap_{p \in \mathcal{B}_{ab}} \mathrm{conv}\left(p, \Delta_{a \succ b}\right),$$

We see that $\Delta_{a \succ b} \subset C$ for every maximal convex set $C \subset \mathcal{B}_{ab}$. Theorem (3.3) then completes the proof.

($\Leftarrow$) If $\mathcal{B}_{ab}$ is star-shaped with respect to every element of $\Delta_{a>b}$, then since $\delta_s \in \Delta_{a>b}$ for every $\delta_s \in \mathcal{B}_{ab}$, condition (s-B$\succ$) is satisfied. $\square$

So far we have only studied the properties of $\mathcal{B}_{ab}$ when preferences satisfy the inter-context conditions (s-B$\succ$) and (WP). If in addition preferences satisfy (Asy.) and (C'ty), as defined in section (3.2), we arrive at the following non-trivial property of context preferences.

Proposition 3.3. *Conditions (Asy.), (C'ty), (s-B$\succ$) and (WP) together imply that for every $a, b \in A$, the set $\mathcal{N}_{ab}$ is a connected separator.*

Although the property of connectedness may at first seem like a purely mathematical property of preferences that has little bearing on behaviour, it is in fact an extremely useful and often intuitive property for context preferences to have. Like continuity and convexity, connectedness is a property of the way preferences vary with p; it is a minimal property of "togetherness" of the direction of preference across contexts.

Connectedness is of course much weaker than convexity, and whilst convexity in itself may not belong in a list of rationality conditions (Binmore 2009 p162), it may be broken down into its intuitive and less intuitive parts. As a result of the above results which provide a set of reasonable sufficient conditions for connectedness of both $\mathcal{B}_{ab}$ and $\mathcal{N}_{ab}$, we argue that the intuitive part is connectedness and the less intuitive part is local-convexity. The latter is defined as follows for any real or complex vector space V.

Definition 3.5 (Schoenberg 1941: local convexity (LC))**.** *Let $p \in Z \subset V$. The set Z is said to be* locally convex *at p if there exists O in the topology of V*

containing p and $Z \cap O$ is convex. The set Z is locally convex if it is locally convex at all its points.

Convexity is implied by condition (p-SB) of chapter 2 and is obviously sufficient for connectedness and local-convexity and the following theorem proves the converse for the case where V is a normed vector space.

Theorem 3.4 (Schoenberg 1941, originally Tietze 1928)**.** *A closed and connected set Z in V which is locally convex is also convex.*

We now prove proposition (3.3).

Proof. Throughout this proof, without loss of generality, we fix $a, b \in A$ and assume that $\varnothing \neq \mathcal{B}_{ab} \neq \Delta$. This we can do because $\mathcal{W}_{ab} = \mathcal{B}_{ba}$, and because $\mathcal{N}_{ab}$ is trivially connected whenever it is empty, and when it is equal to Δ it is connected because Δ is a connected set, as indeed is any simplex.

In step (1) we show that there exists a unique maximal connected subset (which called a component and is defined below) of $\Delta \backslash \mathcal{B}_{ab}$, C' that contains the set $\mathcal{W}_{ab} \dot\cup \Delta_{a\sim b}$, where

$$\Delta_{a\sim b} := \operatorname{conv}\left(\{\delta_s : a \sim_s b\}\right).$$

We do so by considering the cases where $\mathcal{W}_{ab}$ is empty and nonempty separately. The latter is a particularly intricate step, for which we introduce some concepts and theorems from topology.

In step (2), we show that $\Delta \backslash C'$ is a subset of $\mathcal{B}_{ab}$, so that in fact $\mathcal{B}_{ab} \dot\cup C' = \Delta$.

This argument is an augmented version of the proof by induction of proposition (3.2). It is significantly complicated by the fact that $\Delta\backslash C'$ is not a simplex. This step completes the proof for the case where $\mathcal{W}_{ab}$ is empty, for then $C' = \mathcal{N}_{ab}$ which, by construction, a connected set.

Finally, given the symmetry between $\mathcal{W}_{ab}$ and $\mathcal{B}_{ab}$, we know that there exists a maximal connected subset of $\Delta\backslash\mathcal{W}_{ab}$, C'' with the following properties: by step (1) $\mathcal{B}_{ab}\dot{\cup}\Delta_{a\sim b} \subset C''$; and by step (2), $\mathcal{W}_{ab}\dot{\cup}C'' = \Delta$. This implies that

$$\mathcal{N}_{ab} := \Delta\backslash\big(\mathcal{B}_{ab}\dot{\cup}\mathcal{W}_{ab}\big) = C' \cap C''$$

contains the component subset of $\mathcal{N}_{ab}$ that contains $\Delta_{a\sim b}$. The proof is then completed by the following lemma.

Lemma 3.1. *Under the conditions of proposition (3.3) suppose that $\mathcal{B}_{ab} \neq \varnothing \neq \mathcal{W}_{ab}$. If C' is the component of $\Delta\backslash\mathcal{B}_{ab}$ containing $\mathcal{W}_{ab}$ and C'' is the component of $\Delta\backslash\mathcal{W}_{ab}$ containing $\mathcal{B}_{ab}$, then $C' \cap C''$ is a connected subset of $\mathcal{N}_{ab}$.*

Proof. As $\mathcal{W}_{ab} \subset C'$ and C' is closed, $\mathrm{cl}_\Delta \mathcal{W}_{ab}$ is a subset of C'. Since $\Theta'' := \mathrm{bd}_\Delta C'' \subset \mathrm{bd}_\Delta(\Delta\backslash\mathcal{W}_{ab}) = \mathrm{bd}_\Delta \mathcal{W}_{ab}$, we have $\Theta'' \subset C'$. We consider the cases where $\mathrm{Int}_\Delta C'' \cap C'$ is empty and nonempty separately.

If $\mathrm{Int}_\Delta C'' \cap C'$ is empty, then $C'' \cap C' = \mathrm{bd}_\Delta C'' \cap C' = \Theta''$, and, by proposition (3.6) in step (1) below, $C'' \cap C'$ is connected. If not, then as we have shown in step (2), $\Delta = \mathcal{W}_{ab} \cup C''$, and so Θ'' separates $\mathcal{W}_{ab}$ from $\mathrm{Int}_\Delta C''$ in

Δ. Similarly, since

$$\begin{aligned} C' &= C' \cap (\mathcal{W}_{ab} \,\dot\cup\, C'') \\ &= \mathcal{W}_{ab} \,\dot\cup\, (C' \cap C'') \\ &= \mathcal{W}_{ab} \dot\cup\, \Theta'' \,\dot\cup\, (C' \cap \operatorname{Int}_\Delta C''), \end{aligned}$$

Θ'' separates $\mathcal{W}_{ab}$ from $\operatorname{Int}_\Delta C'' \cap C'$. Then because each of C' and Θ'' are connected sets, and $\mathcal{W}_{ab}$ is a component of $C'\backslash\Theta''$, by theorem 5 of Kuratowski p140, $C'\backslash\mathcal{W}_{ab} = C' \cap C''$ is connected. $\square$

Step 1. We first state the definition of a component and prove that it is maximal and hence unique.

Definition 3.6 (cf. Munkres (2000) p175)**.** *Define an equivalence relation $\mathcal{E}$ on Z, such that for all x, y in Z, $x \,\mathcal{E}\, y$ if and only if there does not exist a separation U, V of Z with $x \in U$ and $y \in V$. The equivalence classes of $\mathcal{E}$ define the (connected)* components *of Z.*

Remark 3.3. *Let $X \subset Z$, where X is a connected set. Consider the union of all connected subsets of Z that contain X, that is*

$$\widehat{C} := \bigcup_\alpha \{C_\alpha \subset Z : X \subset C_\alpha,\ C_\alpha \text{ is connected}\}.$$

As $\{C_\alpha\}$ is a collection of connected subspaces of Z with X in common, theorem 23.3 of Munkres (2000) implies that $\widehat{C}$ forms a connected set. This implies that there exists no separation of this union, and so $\widehat{C}$ is contained in one of the equivalence classes of the relation $\mathcal{E}$ in definition (3.6). Thus, there exists a component C^* of Z that contains $\widehat{C}$.

On the other hand, as C^* is connected and it contains X, it must be a member of the collection $\{C_\alpha\}$, so that $C^* \subset \widehat{C}$. This shows that once we have identified a connected subset X of a set Z, the component containing X is uniquely defined to be the largest set containing X. We will make much use of this property in what follows.

We will also need the following proposition regarding the components of a closed set.

Proposition 3.4. *Let K be a closed subset of Z, and let O and P be a separation of K. Then O and P are closed in Z. Equivalently, $Z\backslash O$ and $Z\backslash P$ are elements of the topology $\mathcal{T}_Z$.*

Proof of proposition 3.4. As O and P define a separation of K, O and P are both open and closed in K with respect to the subspace topology

$$\mathcal{T}_K := \{K \cap E : E \in \mathcal{T}_Z\}.$$

Take the closure of O in Z, $\bar{O} := O \cup O'$, where O' is the set of limit points of O in Z. Now since O is closed in K, $O = \bar{O} \cap K$, and if O is not closed in Z there exists a limit point p of O with $p \in \bar{O}\backslash K$. However since $O \subset K$ and $p \notin K$, by the definition of a limit point, every element of $\mathcal{T}_Z$ which contains p has nonempty intersection with K, so p is a limit point of K in the space Z: a contradiction of the fact that K is closed in Z. Thus O is closed in Z.

Since O and P form a separation of K, neither contains any of the limit points in Z of the other. Then as $K = O \cup P$ is closed in Z, $K\backslash O$ contains all its limit points in Z, and so P is closed. $\square$

Lemma 3.2. *Under the conditions of proposition (3.3) let* $\mathcal{W}_{ab} = \varnothing \neq \mathcal{B}_{ab} \neq \Delta$. *Then the simplex* $\Delta_{a\sim b} := \operatorname{conv}\left(\{\delta_s : \delta_s \in \mathcal{N}_{ab}\}\right)$ *is a nonempty connected subset of* $\mathcal{N}_{ab}$.

Proof. First note that $\mathcal{N}_{ab}$ is nonempty because by assumption $\Delta \backslash \mathcal{B}_{ab}$ is nonempty and a subset of $\mathcal{N}_{ab}$. Suppose by way of contradiction that $\mathcal{N}_{ab}$ contains no extremal points $\delta_.$ of Δ. Then since $\mathcal{W}_{ab}$ is empty, the set $\{\delta_s : \delta_s \in \mathcal{B}_{ab}\}$ is a subset of $\mathcal{B}_{ab}$. The induction proof of proposition (3.1) then implies that $\Delta \subset \mathcal{B}_{ab}$ which is the desired contradiction. So the set $\Delta_{a\sim b}$ is nonempty.

Now if $p \in \Delta_{a\sim b}$ then by definition, for all $s \in \operatorname{supp}(p)$ we have $\delta_s \in \mathcal{N}_{ab}$, that is $\neg(a \succ_s b)$. Then by condition (WP), we see that $\neg(a \succ_p b)$, and because $\mathcal{W}_{ab}$ is empty $p \in \mathcal{N}_{ab}$. Moreover, as $\Delta_{a\sim b}$ is a simplex in its own right, it is connected. $\square$

Now by remark (3.3), in the case where $\mathcal{W}_{ab}$ is empty, we may directly define C' to be the component of $\Delta \backslash \mathcal{B}_{ab}$ that contains $\Delta_{a\sim b} \cup \mathcal{W}_{ab}$.

Now for the case where $\mathcal{W}_{ab}$ is nonempty. If both $\mathcal{B}_{ab}$ and $\mathcal{W}_{ab}$ are nonempty, we first show that $\mathcal{B}_{ab}$ and $\mathcal{W}_{ab}$ are separated or disconnected. The following definitions introduce some useful terminology that is used in other chapters of this thesis.

Definition 3.7 (Kuratowski (1966) V.1 p58)**.** *X and Y, both nonempty subsets of Z, are said to be* separated *or* disconnected *if*

$$\bar{X} \cap Y = \varnothing = X \cap \bar{Y}.$$

Definition 3.8. *Let X, Y be subsets of Z. Suppose that for all separations O, P of Z either $X \not\subset O$ or $Y \not\subset P$. If, for some subset Γ of Z, there exists a separation, O', P' of $Z \backslash \Gamma$ with $X \subset O'$ and $Y \subset P'$, then Γ is said to* separate *or* disconnect *X from Y (as well as each $Q \subset X$ from $R \subset Y$). Equivalently, X and Y are said to be separated by Γ.*

Proposition 3.5. *Let X and Y both be nonempty, disjoint, open and subsets of the same component of Z. Then*

(i) X and Y are separated sets,

(ii) $Z \backslash (X \cup Y)$ separates X from Y.

(iii) For arbitrary $x \in X$ and $y \in Y$, if there exists a path $P(x, y)$ in Z between x and y, then $P(x, y)$ has nonempty intersection with $Z \backslash (X \cup Y)$.

Proof. (i) Since both X and Y are disjoint and open in Z, Y is a subset of the set $Z \backslash X$ which, as the complement of an open subset of Z, is closed in Z. Then since the closure of Y in Z is the smallest closed set containing Y, it is contained in $Z \backslash X$. Thus $X \cap \mathrm{cl}_Z Y = \varnothing$. Similarly, $Y \cap \mathrm{cl}_Z X = \varnothing$. Therefore, by definition (3.7), X and Y are separated.

(ii) The fact that X and Y are separated subsets of Z implies that neither contains limit points of the other. This, by lemma 23.1 of Munkres (2000) together with the fact that they are both nonempty, implies that they define a separation of $X \cup Y$. Then since X and Y lie in the same component of Z, $Z \backslash (X \cup Y)$ separates X from Y.

(iii) Suppose there exists $x \in X$ and $y \in Y$ such that a path $P(x, y)$ between them lies in Z. Define K_x and K_y to be the components of X and Y that contain x and y respectively. Now, contrary to part (iii) of this proposition, suppose in addition that $P(x, y) \subset X \cup Y$.

By definition 3.3, since such a path is homeomorphic to a closed interval in $\mathcal{R}$, it is connected. Then then since it contains x and K_x is a component, remark (3.3) implies $P(x, y)$ is in fact a subset of K_x. By the same argument $P(x, y)$ is also a subset of K_y, so that the definition of component implies that K_x is in fact equal to K_y. If this true however, then $X \cap Y$ is nonempty and we have derived a contradiction of the assumption that X and Y are disjoint. $\square$

By virtue of the asymmetry condition on preferences, $\mathcal{B}_{ab}$ and $\mathcal{W}_{ab}$ are both disjoint, and by the continuity condition on preferences, they are both open. Moreover, we are considering the case where they are both nonempty. If we take $\mathcal{B}_{ab}$, $\mathcal{W}_{ab}$ and the connected set Δ to be X, Y and Z respectively, then proposition (3.5) shows that $\mathcal{B}_{ab}$ and $\mathcal{W}_{ab}$ are separated sets, that $\mathcal{N}_{ab} = \Delta \backslash (\mathcal{B}_{ab} \dot{\cup} \mathcal{W}_{ab})$ separates $\mathcal{B}_{ab}$ from $\mathcal{W}_{ab}$.

In fact proposition (3.5) implies that $\mathcal{N}_{ab}$ is nonempty whenever both $\mathcal{B}_{ab}$ and $\mathcal{W}_{ab}$ are nonempty. We will now show that there exists a component of $\mathcal{N}_{ab}$ that separates $\mathcal{B}_{ab}$ from $\mathcal{W}_{ab}$. We do this using a number of results and definitions from Kuratowski (1966).

Definition 3.9. *Two continuous functions $f_0 : X \longrightarrow Y$ and $f_1 : X \longrightarrow Y$ are said to be homotopic (with respect to Y), if there exists a continuous function* *of two variables* $h : X \times [0,1] \longrightarrow Y$ *such that*

$$h(x,0) = f_0(x) \quad \textit{and} \quad h(x,1) = f_1(x). \tag{3.1}$$

Definition 3.10 (Based on Kuratowski (1966) V.2 p370)**.** *The space* X *is said to be* contractible *with respect to the space* Y *if every continuous function* $f : X \longrightarrow Y$ *is homotopic to a constant. That is homotopic to some function* $g : X \longrightarrow Y$ *with* $g(x) = c$ *for all* $x \in X$*.*
A space X *which is contractible with respect to itself is said to be contractible in itself.*

Definition 3.11. *Let* $X \subset Y$*. If the continuous function* $f : X \longrightarrow Y$ *is homotopic to the identity, i.e. if there exists a continuous function* $h : X \times [0,1] \longrightarrow Y$ *such that*

$$h(x,0) = x \quad \textit{and} \quad h(x,1) = f(x), \tag{3.2}$$

the set $f(X)$ *is said to be obtained from* X *by a* deformation *in* Y *(namely* h *is that deformation). If* $f(X) = c$ *for some* $c \in Y$ *then* Y *is said to be deformable to one point.*

By theorem 2 of Kuratowski (1966) V.2 p374 a space is contractible in itself if and only if it is deformable to one point. In this way, the following claim shows that Δ is contractible in itself.

Claim 3.3. *A convex set* Z *is deformable to one point.*

Proof. Choose an arbitrary point c in Z. Since Z is convex, for each x the straight-line paths $[c, x]$ are all contained in Z. Then define $h : Z \times [0,1] \longrightarrow Z$, such that $h(x, t)$ takes values $tc + (1 - t)x$ for each $(x, t) \in Z \times [0, 1]$. This function is clearly continuous in both x and t, and furthermore

$$h(x, 0) = x \quad \textit{and} \quad h(x, 1) = c. \tag{3.3}$$

By definition 3.11 therefore, Z is deformable to one point. □

We are now ready to prove that every component of the complement in Δ of the open, connected set $\mathcal{B}_{ab}$ has a connected boundary.

Proposition 3.6. *Let Z be contractible in itself and Y be a connected set. If X is a component of $Z \backslash Y$, then* $\mathrm{bd}_Z X$ *is connected.*

Proof. This is a direct consequence of theorem 6 of Ch.VIII section 57, II on p437 together with theorem 9(i) of section 57, I on p435 both in Ch.VIII of Kuratowski (1966) Vol.2. □

Since $\mathcal{W}_{ab}$ is connected, nonempty, and contained in the complement in Δ of the connected subset $\mathcal{B}_{ab}$, it is contained in some component of $\Delta \backslash \mathcal{B}_{ab}$. Let this component be denoted by C'. Then by proposition (3.6), the boundary of C' in Δ is a connected set which we denote by Θ'.

Definition 3.12. *Let Ξ denote the component of $\mathcal{N}_{ab}$ that contains Θ'.*

Claim 3.4. *If $\mathcal{B}_{ab} \neq \varnothing \neq \mathcal{W}_{ab}$ and Ξ is defined as in definition (3.12), then for every $t \in S$ such that $a \sim_t b$, δ_t lies in Ξ.*

Proof. By the assumption that both $\mathcal{B}_{ab}$ and $\mathcal{W}_{ab}$ are nonempty, open and disjoint we know, by proposition (3.5), that $\mathcal{N}_{ab}$ is nonempty. In the case where the set $\{t \in S : a \sim_t b\}$ is empty, the claim is trivially true, so without loss of generality, suppose this set is nonempty.

Recall that $\mathcal{B}_{ab} \neq \varnothing \neq \mathcal{W}_{ab}$ together with condition (WP) implies that both the sets $\{s \in S : a \succ_s b\}$ and $\{u \in S : b \succ_u a\}$ are nonempty. So consider any triple $s, t, u \in S$ satisfying $\delta_s \in \mathcal{B}_{ab}$, $\delta_t \in \mathcal{N}_{ab}$ and $\delta_u \in \mathcal{W}_{ab}$.

Take the path $P(s,t,u) := \Delta(s,t) \cup \Delta(t,u)$ that runs along the boundary of the 2-simplex $\Delta(s,t,u)$ from δ_s to δ_u via δ_t. Recall that $\Theta := \mathrm{bd}_\Delta\, C$ separates $\mathrm{Int}_\Delta\, C$ from $\Delta \backslash C$. The fact that $\delta_s \in \mathcal{B}_{ab} \subset \Delta \backslash C$ and $\delta_u \in \mathcal{W}_{ab} \subset \mathrm{Int}_\Delta\, C$ in conjunction with part (iii) of proposition (3.5), implies that the path $P(s,t,u)$ has nonempty intersection with Θ.

Since Ξ is the component of $\mathcal{N}_{ab}$ containing Θ, the proof will be complete if we show that $\mathcal{N}_{ab} \cap P(s,t,u)$ is connected, for this set will then contain both δ_t and $\Theta \cap P(s,t,u)$. To this end we define a homeomorphism $f : [0,1] \longrightarrow P(s,t,u)$ such that $f(0) = \delta_s$ and $f(1) = \delta_u$, and an order $\leqslant_f$ on $P(s,t,u)$ such that for all $x, y \in [0,1]$

$$x \leqslant y \quad \Leftrightarrow \quad f(x) \leqslant_f f(y).$$

Define q to be the l.u.b. of $\mathcal{B}_{ab} \cap \Delta(s,t)$. Then $\delta_s <_f q$ by virtue of the fact that $\mathcal{B}_{ab} \cap \Delta(s,t)$ contains δ_s and is open in $\Delta(s,t)$. As $\delta_t \in \mathcal{N}_{ab}$ and $\delta_u \in \mathcal{W}_{ab}$, condition (WP) implies that $\mathcal{B}_{ab} \cap \Delta(t,u)$ is empty. Thus

$$\mathcal{B}_{ab} \cap P(s,t,u) = \mathcal{B}_{ab} \cap \Delta(s,t),$$

so that δ_t is an upper bound of $\mathcal{B}_{ab} \cap P(s,t,u)$ and $q \leqslant_f \delta_t$.

Take $p \in \mathcal{B}_{ab} \cap \Delta(s,t)$ arbitrarily close to q. Then since $p \in \mathcal{B}_{ab}$, condition (s-B$\succ$) implies that all points r satisfying $r <_f p$ lie in $\mathcal{B}_{ab}$. Thus if $p <_f q$ then $p \in \mathcal{B}_{ab}$. Thus $\mathcal{B}_{ab} \cap P(s,t,u)$ is a (connected) segment of $\Delta(s,t)$.

By an identical argument to the previous paragraph conditions (WP) and (s-B$\succ$) together imply that $\mathcal{W}_{ab} \cap P(s,t,u)$ is a (connected) segment of $\Delta(t,u)$ with a g.l.b. q' such that $q' <_f p$ implies $p \in \mathcal{W}_{ab}$. Then the definition of l.u.b. and g.l.b. implies that every element p satisfying $q <_f p <_f q'$ lies in neither $\mathcal{B}_{ab}$ nor $\mathcal{W}_{ab}$: thus $p \in \mathcal{N}_{ab}$. Finally, the points q and q' also lie in $\mathcal{N}_{ab}$ because $\mathcal{N}_{ab}$ is closed, and therefore

$$P(s,t,u)\backslash(\mathcal{B}_{ab} \cup \mathcal{W}_{ab}) \equiv P(s,t,u) \cap \mathcal{N}_{ab}$$

is equal to the (connected) path $P(q,q') \subset P(s,t,u)$. $\square$

As a result of this claim and lemma (3.2), we know that $\Delta_{a\sim b}$ is a subset of the component Ξ of $\mathcal{N}_{ab}$ that contains the boundary of C', Θ'.

To complete step (1) of this proof it only remains to be shown that Ξ is a subset of C'. This follows because Θ' lies in the intersection of Ξ and C', both of which lie in the set $\Delta\backslash\mathcal{B}_{ab}$. This implies, via theorem 23.3 of Munkres p.166, that $\Xi \cup C'$ is a connected subset of $\Delta\backslash\mathcal{B}_{ab}$, so that $\Xi \subset C'$ because C' is a component of $\Delta\backslash\mathcal{B}_{ab}$. We summarize step (1) by the following lemma.

Lemma 3.3. *Under the conditions of proposition (3.3); for each $a, b \in A$ there exists a component C' of the complement of $\mathcal{B}_{ab}$ in Δ such that*

$$\mathcal{W}_{ab} \dot{\cup} \Delta_{a\sim b} \subset C'.$$

Step (2) We recall that the purpose of step (2) is to prove that $\mathcal{B}_{ab}\dot{\cup}C' = \Delta$, or, equivalently, that

$$\Delta\backslash C' \subset \mathcal{B}_{ab}.$$

To reduce notational clutter, in the remainder of the proof we simply refer to the set C' as C.

Definition 3.13. *Let C denote the component of $\Delta\backslash\mathcal{B}_{ab}$ that contains $\Delta_{a\sim b} \cup \mathcal{W}_{ab}$ and define D to be its complement in $\Delta\backslash\mathcal{B}_{ab}$. That is*

$$D := (\Delta\backslash\mathcal{B}_{ab})\backslash C.$$

Lemma 3.4. *D is equal to $\mathcal{N}_{ab}\backslash C$, and if D is nonempty then $\mathcal{N}_{ab} \cap C$ and D form a separation of $\mathcal{N}_{ab}$.*

Proof of lemma (3.4). Note that since neither C nor D contain elements of $\mathcal{B}_{ab}$, and C contains $\mathcal{W}_{ab}$, any element of D must lie in $\mathcal{N}_{ab}\backslash C$. Conversely, if p lies in $\mathcal{N}_{ab}\backslash C$, then since it is an element of $\mathcal{N}_{ab}$ it lies in $\Delta\backslash\mathcal{B}_{ab}$, and since it does not lie in C, it must lie in D. Thus $D = \mathcal{N}_{ab}\backslash C$.

Suppose D is nonempty. We first show that C and D form a separation of $\Delta\backslash\mathcal{B}_{ab}$. Since C and D are by definition disjoint with union equal to the whole of $\Delta\backslash\mathcal{B}_{ab}$, we only need to show that they are open in $\Delta\backslash\mathcal{B}_{ab}$. Since C is a component of $\Delta\backslash\mathcal{B}_{ab}$ it is both open and closed in this set. In this case, its complement in $\Delta\backslash\mathcal{B}_{ab}$, D, is also both open and closed in $\Delta\backslash\mathcal{B}_{ab}$.

In the case that D is nonempty, the preceding paragraph shows that C and D form a separation of $\Delta\backslash\mathcal{B}_{ab}$. Now since Δ is a connected set, by definition

(3.8), $\mathcal{B}_{ab}$ separates C from D. By the same definition, $\mathcal{B}_{ab}$ also separates $\mathcal{N}_{ab} \cap C$ from $\mathcal{N}_{ab} \cap D = D$. Thus $\mathcal{N}_{ab} \cap C$ and D are separated sets whose union is, by the first part of this lemma, equal to $\mathcal{N}_{ab}$. That is $\mathcal{N}_{ab} \cap C$ and D form a separation of $\mathcal{N}_{ab}$. □

Lemma 3.5. *If $\emptyset \neq \mathcal{B}_{ab} \neq \Delta$, $\Delta_{a>b} \equiv \{p \in \Delta : \forall s \in \mathrm{supp}(p),\ a \succ_s b\}$, and C and D are defined as in definition 3.13, then there exists an open subset U of Δ, with the following properties:*

(1) $C \subset U$, and

(2) $\widetilde{D} := D \cup \Delta_{a>b}$ is contained in the complement of U in Δ.

Proof of lemma 3.5. By lemma (3.4), both C and D are closed. As $\Delta_{a>b}$ is a subsimplex of Δ, it contains its relative boundary and is a subset of Δ, so that it is closed in Δ. By proposition (3.2) $\Delta_{a>b}$ is also contained in $\mathcal{B}_{ab}$, and so it is disjoint from $C \subset \Delta \backslash \mathcal{B}_{ab}$. Hence, $\widetilde{D} := D \dot{\cup} \Delta_{a>b}$, as the union of two closed sets, is also closed and, moreover, $\widetilde{D}$ has empty intersection with C.

By Theorem 32.3 of Munkres (2000) p218, every compact Hausdorff space is normal, and thus Δ is a normal space. This means means that there exists a pair of disjoint open sets $U, \widetilde{U}$ in Δ such that $C \subset U$ and $\widetilde{D} \subset \widetilde{U}$. Then, since $\widetilde{U} \subset \Delta \backslash U$, $\widetilde{D} \subset \Delta \backslash U$.

□

Lemma 3.6. *Suppose $\emptyset \neq \mathcal{B}_{ab} \neq \Delta$; C, D are defined as in definition 3.13 and $\widetilde{D} := D \cup \Delta_{a>b}$. Take U to be a set satisfying the properties of lemma 3.5, then there exists a set, which we denote by V, with the following properties:*

(1) V is open, contains C and its closure in Δ, $\bar{V}$, is a subset of U;

(2) the boundary of V in Δ, $\mathrm{bd}_\Delta V$, is a subset of $\mathcal{B}_{ab}$, and

(3) $\widetilde{D}$ is a subset of the interior of $\Delta \backslash V$ relative to Δ.

In claim (iii), we have shown that C is closed in Δ, so that Lemma 31.1 (b) of Munkres on p212 establishes (1). We now prove (2) and (3) for any set V with property (1).

Proof of lemma 3.6. The complement of C in $\Delta \backslash \mathcal{B}_{ab}$, which have denoted by D, is a subset of $\Delta \backslash U$ by lemma (3.5). By the same lemma, and the fact that $\mathcal{W}_{ab}$ is contained in C, $U \backslash C$ is a subset of $\mathcal{B}_{ab}$. This implies that $\bar{V}$, as a subset of U, contains only elements of C and $\mathcal{B}_{ab}$. Then by part (1) of this claim V is open in Δ, so that $\mathrm{bd}_\Delta V$ is a subset of $\Delta \backslash V$. Finally, as $C \subset V$, the intersection of C and $\mathrm{bd}_\Delta V$ is empty, and (2) is proven.

To prove (3) we first note that by lemma (3.5) $\widetilde{D}$ is a subset of $\Delta \backslash U$, and by part (1) of this lemma $\bar{V}$ is a subset of the open set U, thus $\Delta \backslash \bar{V} = \mathrm{Int}(\Delta \backslash V)$ contains $\Delta \backslash U$ and the proof is complete. $\square$

We will now show that $Q := \Delta \backslash V$ is a subset of $\mathcal{B}_{ab}$. Similar to the proof of proposition (3.2), we do this by induction on the intersection of the unit subsimplices of Δ with Q. That is, we take Δ_k to be an arbitrary k-subsimplex of Δ, and we define

$$Q_k := Q \cap \Delta_k \equiv \Delta_k \backslash V, \tag{3.4}$$

then the basis of the argument is that once we know that the relative boundary of Q_k is contained in $\mathcal{B}_{ab}$ then we can apply condition (s-B$\succ$) together with claim (3.6) in the appendix to show that the whole of Q_k lies in $\mathcal{B}_{ab}$. To this end, the following lemma provides a useful decomposition of the relative boundary.

Lemma 3.7. *If $X \subset Y \subset Z$ and Y is closed in Z, then the closure of X in Y is equal to the closure of X in Z, and furthermore*

$$\mathrm{bd}_Z\, X = \mathrm{bd}_Y\, X \cup \big(X \cap \mathrm{bd}_Z\, Y\big). \tag{3.5}$$

Proof of lemma 3.7. To see that $\mathrm{cl}_Y(X) = \mathrm{cl}_Z(X)$, we first note that $\mathrm{cl}_Y(X) \subset \mathrm{cl}_Z(X)$ simply because Y is a subset of Z. For the reverse inclusion, we note that all the limit points of X in Z are contained in $\mathrm{cl}_Y(X)$ because Y is closed in Z and it contains X.

To show that equation (3.5) holds, we first note that by the definition of the boundary and basic set theory we have

$$\begin{aligned}
\mathrm{bd}_Z\, X &:= X \cap \mathrm{cl}_Z(Z \backslash X) \\
&= X \cap \mathrm{cl}_Z(Y \cup Z \backslash Y) \backslash X) \\
&= X \cap \mathrm{cl}_Z(Y \backslash X \cup (Z \backslash Y) \backslash X)
\end{aligned}$$

where $(Z\backslash Y)\backslash X = Z\backslash Y$ as $X \subset Y$. Then since the closure of a finite union is equal to the finite union of the closures,

$$\mathrm{bd}_Z\, X = \big(X \cap \mathrm{cl}_Z(Y \backslash X)\big) \cup \big(X \cap \mathrm{cl}_Z(Z \backslash Y)\big) \tag{3.6}$$

Then since $Y\backslash X \subset Y$ we may apply the result of the first part of this lemma to the first term of the union in equation 3.6, and use the fact that $X = X \cap Y$

to complete the proof. That is

$$\operatorname{bd}_Z X = \big(X \cap \operatorname{cl}_Y(Y\backslash X)\big) \cup \big(X \cap (Y \cap \operatorname{cl}_Z(Z\backslash Y)\big),$$

where the first term in the union is equal to $\operatorname{bd}_Y X$ and the second term is equal to $\operatorname{bd}_Z \big(X \cap \operatorname{bd}_Z Y\big)$. □

Part (3) of lemma (3.6) shows that Q has nonempty interior relative to Δ, so that the affine hull of Δ and Q are the same. A similar statement can be made for all the k-subsimplices of Δ that have nonempty intersection with the interior of Q. These are precisely the subsimplices that are either subsets of $\Delta_{a \succ b}$, or that have at least one extremal element δ_s in $\mathcal{B}_{ab}$ and at least one other δ_t in $\Delta\backslash\mathcal{B}_{ab} = \mathcal{N}_{ab} \cup \mathcal{W}_{ab}$. All other subsimplices lie in $\Delta_{a\sim b} \cup \Delta_{b\succ a}$, and are therefore wholly contained in C, so that they do not intersect Q. Let Δ_k be an arbitrary k-subsimplex of Δ and endow Δ_k with the subspace topology, then $Q_k := Q \cap \Delta_k$ has the same affine hull as Δ_k whenever $\varnothing \neq Q_k$. We denote this space by K, for in this case the relative interiors of both Q_k and Δ_k are homeomorphic to a k-dimensional open subset of $\mathcal{R}^k$.

Since $Q_k \subset \Delta_k \subset K$, lemma (3.7) shows that the boundary of Q_k in K can be decomposed into the following union where bd_k denotes the boundary of a set with respect to K:

$$\operatorname{bd}_K Q_k = \operatorname{bd}_{\Delta_k} Q_k \cup \big(Q_k \cap \operatorname{bd}_K \Delta_k\big). \tag{3.7}$$

The next lemma shows that the first term in the above union lies in $\mathcal{B}_{ab}$.

Lemma 3.8. *Let Q_k be defined as in equation (3.4). Then $\operatorname{bd}_{\Delta_k} Q_k$ is a subset* *of* $\mathcal{B}_{ab}$.

Proof. Recall that $Q_k := Q \cap \Delta_k$. By part (1) of lemma (3.6) $Q := \Delta \backslash V$ is the complement of an open set in Δ, and as such it is closed in Δ. Then since Δ_k is endowed with the subspace topology, $Q \cap \Delta_k$ is closed in Δ_k. Thus

$$\begin{aligned}\operatorname{bd}_{\Delta_k}(Q_k) &:= \big(Q \cap \Delta_k\big) \cap \operatorname{cl}_{\Delta_k}\big(\Delta_k \backslash (Q \cap \Delta_k)\big) \\ &= \big(Q \cap \Delta_k\big) \cap \operatorname{cl}_{\Delta_k}\big((\Delta_k \backslash Q) \cup \varnothing\big) \end{aligned} \tag{3.8}$$

Then since $\Delta_k \cap Q \subset \Delta_k \subset \Delta$ and Δ_k is closed in Δ, lemma (3.7) implies that

$$\operatorname{cl}_{\Delta_k}(\Delta_k \backslash Q) = \operatorname{cl}_{\Delta}(\Delta_k \backslash Q).$$

Then substituting Δ for Δ_k in equation (3.8), we obtain the following inclusion:

$$\operatorname{bd}_{\Delta_k}(Q_k) \subset \big(Q \cap \Delta\big) \cap \operatorname{cl}_{\Delta}\big(\Delta \backslash Q\big),$$

where the right-hand-side is of course the definition of $\operatorname{bd}_{\Delta} Q$. Then by lemma (3.6), this set is contained in $\mathcal{B}_{ab}$. $\square$

It remains to be shown that the second term of the union in equation (3.7), $Q_k \cap \operatorname{bd}_k \Delta_k$, lies in $\mathcal{B}_{ab}$. First we note that because Δ_k is closed in K, $\operatorname{bd}_k \Delta_k$ is a subset of Δ_k, and so

$$Q_k \cap \operatorname{bd}_k \Delta_k := (Q \cap \Delta_k) \cap \operatorname{bd}_k \Delta_k = Q \cap \operatorname{bd}_k \Delta_k.$$

We will show this by an induction argument on the subsimplices of Δ.

We note that because the boundary of a k-dimensional unit simplex is the union of its (k–1)-dimensional faces, we have

$$Q_k \cap \mathrm{bd}_k\, \Delta_k = \bigcup_l \big\{ Q \cap \Delta_l : \, l = k-1, \, \Delta_l \subset \Delta_k \big\}.$$

As all the 0-simplices in Q are contained in $\mathcal{B}_{ab}$, if $Q \cap \Delta_k = \Delta_k$, then $Q_k \cap \mathrm{bd}_k\, \Delta_k$ contains all its extremal measures, and by condition (s-B$\succ$) on preferences it is a subset of $\mathcal{B}_{ab}$. If on the other hand $Q \cap \Delta_k \neq \Delta_k$, then Δ_k has at least one 0-simplex in $\mathcal{B}_{ab}$ and at least one other in $\Delta \backslash \mathcal{B}_{ab}$.

Consider an arbitrary 1-subsimplex, $\Delta(s,t)$, of Δ with $\delta_s \in \mathcal{B}_{ab}$ and $\delta_t \in \Delta \backslash \mathcal{B}_{ab}$. $\Delta(s,t)$ is a compact convex subset of its 1-dimensional affine hull, and δ_s is an extremal point of $\Delta(s,t)$, which we treat as the point u of claims (3.5), (3.6) and (3.7) in the appendix to this chapter. Under the induced order $\leqslant_f$ on $\Delta(s,t)$, δ_s and δ_t are defined to be the g.l.b and l.u.b. of $\Delta(s,t)$, respectively. Moreover, as Q is closed, $Q \cap \Delta(s,t)$ is a closed subset of $\Delta(s,t)$, and the l.u.b. $Q \cap \Delta(s,t)$, which we denote by p satisfies $p <_f \delta_t$ as $\delta_t \in V$. Thus, by claim 3.7 of the appendix to this chapter, $p \in \mathrm{bd}_\Delta\, Q \subset \mathcal{B}_{ab}$. So that in fact every point of $Q \cap \Delta(s,t)$ is an element of $\mathcal{B}_{ab}$.

The proof of the general case is completed via the induction hypothesis. That is, suppose the intersection of every (k-1)-subsimplex of Δ with Q lies in $\mathcal{B}_{ab}$. Then for every k-subsimplex of Δ, Δ_k, $Q_k \cap \mathrm{bd}_k\, \Delta_k \subset \mathcal{B}_{ab}$, so that the whole of the relative boundary of $Q \cap \Delta_k$ is contained in $\mathcal{B}_{ab}$. Then by claim (3.6) of the appendix to this chapter and condition (s-B$\succ$), the whole of $Q \cap \Delta_k$ is a subset of $\mathcal{B}_{ab}$.

We have therefore shown that $Q \cap \Delta = Q \subset \mathcal{B}_{ab}$, so that D has empty intersection with $Q = \Delta \backslash V$. Recalling that $V \cap D$ is also empty, we see that D is in fact empty, so that $C \cup \mathcal{B}_{ab} = \Delta$ and the only component of $\Delta \backslash \mathcal{B}_{ab}$ is C. $\square$

3.4 Summary of results

The following theorem summarizes the results of this chapter in the form of an equivalence between conditions on preferences and properties of the representation of preferences.

Theorem 3.5. *The following two statements are equivalent.*

1) *Context preferences $\{(A, \succ_p) : p \in \Delta\}$ satisfy the conditions (Asy.), (C'ty), (s-B) and (NT).*

2) *Context preferences have a continuous, ordinal utility representation with the following properties: for all a, b in A*

   i) *$\{p : U(a,p) = U(b,p)\}$ is a connected separator containing $\Delta_{a\sim b}$;*

   ii) *$\{p : U(a,p) > U(b,p)\}$ is star-convex w.r.t. $\Delta_{a\succ b}$, and its complement is star-convex w.r.t $\Delta_{b\succsim a}$.*

Imposing conditions (T) and (LC) (thinness and local convexity of the no-strict-preference relation) would give rise to a choice of representation type. The first being the expected preference representation of chapter (2) and the second being a utility function of the form presented in this chapter which is non-linear in contexts. Going further, and imposing the diversity conditions

(Div.) of that chapter would give rise to an expected utility representation as an alternative to that of the present chapter.

Finally we note that although the results of this and the last section hold for probability simplices of countable dimension, we have not assumed that Δ is a metric space. Instead we have used a more general approach and relied upon results that are contained in the following appendix.

3.5 Appendix

If E is a topological vector space over $\mathcal{R}$ (with the Hausdorff property). Then the affine hull P of any two distinct points v', w' of E is a well defined 1-dimensional subset of E. Indeed we can define a homeomorphism between P and $\mathcal{R}$ in the following way: for all $x \in \mathcal{R}$, let $f(x) = xw' + (1-x)v'$, so that the image of f is indeed the affine hull of v and w, and it is clearly one-to-one with $f(0) = v'$ and $f(1) = w'$. Moreover, if we define the order $\leqslant_f$ such that for all $x, y \in \mathcal{R}$ we have

$$x \leqslant y \quad \Leftrightarrow \quad f(x) \leqslant_f f(y)$$

then f is an order isomorphism between $(\mathcal{R}; \leqslant)$ and$(P; \leqslant_f)$, and as such every subset of P has a least upper bound (lub) and a greatest lower bound (glb) under $\leqslant_f$.

Let Z be a compact, convex subset of E with cardinality greater than two. Define $Q_u := \{P_\alpha := \{x\alpha + (1-x)u : x \in \mathcal{R}\} : \alpha \in Z\}$ to be the collection of affine hulls through pairs of points of Z, such that for every pair, one element of the pair is the point u. We define u to be an extremal point of Z if for every P_α in Q_u, u is the glb of $P_\alpha \cap Z$. Thus, by definition, if u is an extremal point of Z, then for an arbitrary element P_α of Q_u with $f : \mathcal{R} \longrightarrow P_\alpha$, $f(x) := x\alpha + (1-x)u$, then $u \leqslant_f v$ for all v in $P_\alpha \cap Z$.

Claim 3.5. *If u is an extremal point of the compact, convex set with cardinality greater than two, Z, then it lies in the boundary of Z relative to its affine hull (henceforth we refer to this as its relative boundary). Moreover, for every $P \in Q_u$, the lub of $P \cap Z$ lies in the relative boundary of Z, and $P \cap Z$ is order homeomorphic to a compact interval in $\mathcal{R}$.*

Proof. Suppose not. Take an arbitrary element P of Q_u generated by u and α. Then either $u \in P \backslash Z$ or $u \in W := P \cap \mathrm{ri}(Z)$, where $\mathrm{ri}(\cdot)$ denotes the interior of a set relative to its affine hull (its relative interior). Since both P and Z are closed subsets of E, $P \cap Z$ is a closed subset of Z. This implies that $P \cap Z$ is compact (Munkres 2000 p181). Then by the extreme value theorem (Munkres 2000 p190) there exists a points v' and w' in $P \cap Z$ such that $f^{-1}(v') \leqslant f^{-1}(u') \leqslant f^{-1}(w')$ for every $u' \in P \cap Z$. This implies that $v' \leqslant_f u'$ for every $u' \in P \cap Z$, so that $v' = u$ and u is therefore an element of $P \cap Z$. Therefore u is an element of W.

The extreme value theorem also implies that $P \cap Z$ contains its lub which we denote by v. Since both P and Z are convex, so is $P \cap Z$. Thus $P \cap Z$ is a compact convex set, and it therefore contains every point u' such that $u \leqslant_f u' \leqslant_f v$ lies in $P \cap Z$. This implies that $1 = f^{-1}(\alpha) \leqslant f^{-1}(v) =: y$, and by compactness there exists $w \in P$ such that $v \leqslant_f w$ so that $y < \infty$. Thus for all $u' \in P \cap Z$, $0 = f^{-1}(u) < f^{-1}(u') \leqslant f^{-1}(v) = y$. In other words, there exists a compact interval $[0, y]$ that is homeomorphic to $P \cap Z$.

If $v \in W := \mathrm{ri}(Z) \cap P$, since $\mathrm{ri}(Z)$ is open in Z and $P \cap Z$ is a subspace of P, W is open in P and strictly contained in $P \cap Z$. Thus $f^{-1}(W)$ is open in $\mathcal{R}$ and strictly contained in $[0, y]$. If z is the supremum of $f^{-1}(W)$, then z cannot be contained in $f^{-1}(W)$, for there would exist $\epsilon > 0$ such that $z + \epsilon \in f^{-1}(W)$ and $z < z + \epsilon$. Thus $z \notin f^{-1}(W)$, so that $x < z$ for every $x \in f^{-1}(W)$; however if this true $y < z$ so that $z \notin P \cap Z$. The same proof shows that u lies in the relative boundary of Z. ☐

We now make a suitable extension of this result to arbitrary closed subsets of Z.

Claim 3.6. *Let E, Z, Q_u and u be defined as above. If K is a closed subset of Z then every element α of K lies in a set $P \cap Z$ for some $P \in Q_u$, and the extremal point u is a lower bound for the set $P \cap K$. Moreover, the lub of $P \cap K$ is contained in the relative boundary of K.*

Proof. The definition of Q_u is such that every point in Z and therefore *a fortiriori* K, is contained in the intersection of Z with some affine hull $P \in Q$. Fix a point $\alpha \in K$. As K is a closed subset of Z, and the latter is compact, K is also compact. Similarly, $P \cap K$ is compact and by the extreme value theorem, it contains both its glb and its lub, and we denote these by v and w, respectively. Since $v \in P \cap K \subset P \cap Z$, $u \leqslant_f v$ and thus u is a lower bound for $P \cap K$. Then by an identical argument to the preceding claim, we see that w is an element of the relative boundary of K. ☐

We will need a slightly stronger result in what follows.

Claim 3.7. *If* $\sup(P \cap K) \neq \sup(P \cap Z)$*, then the least upper bound of* $P \cap K$ *is an element of* $\operatorname{bd}_Z K$*, the boundary of* K *relative to* Z*.*

The intuition behind this claim is that the relative boundary of K is not in general equal to the boundary of K relative to Z. The case where they are not equal is where K intersects the relative boundary of Z, so that part of the relative boundary of K is contained in the interior of K in the topology of Z.

Proof. If $k := \sup(P \cap K) \neq z := \sup(P \cap Z)$, then $k <_f z$ as K is a subset of Z. By claims (3.5) and (3.6), k is an element of the closed set $P \cap K$ and $z \in (P \backslash K) \cap Z = (P \cap Z) \backslash K)$. Since $P \cap Z$ is isomorphic to $[0, y]$ for some $0 < y < \infty$, with $f^{-1}(z) = y$, if we define $x := f^{-1}(k)$, then $(x, y]$ is open in $[0, y]$ and contained in $W := f^{-1}((P \cap Z) \backslash K)$. Moreover, since $x \in f^{-1}(P \cap K)$, x separates $(x, y]$ from $W \backslash (x, y]$. Then by continuity of f^{-1} and the fact that connectedness is a topological property, $f([x, y]) = \operatorname{cl}_P f((x, y])$ and it is a closed component of $\operatorname{cl}_P((P \cap Z) \backslash K)$. Thus k is an element of the set

$$\begin{aligned} \big((P \cap Z) \cap K\big) \cap \operatorname{cl}_P((P \cap Z) \backslash K)) &\subset \big(Z \cap K\big) \cap \operatorname{cl}_Z(Z \backslash K)) \\ &\equiv \operatorname{bd}_Z K \end{aligned} \tag{3.9}$$

□

Chapter 4

Extending preferences to contexts

4.1 Introduction

When a decision-maker sees no benefit to fooling her opponent in a game, she presumably sees no reason to define her preferences over her own mixed strategies. Yet if the decision-maker is to be modeled using a von Neumann–Morgenstern (1944) (henceforth [vNM]) expected utility function, or one of the generalizations we discuss in more detail below, this is what she is required to do.

There are models, such as that of Gilboa and Schmeidler (2003) [GS] and chapter 2 of this thesis, which address this concern by reducing the domain of preferences to be the set of alternatives (pure actions) that are available to the decision-maker. In this case, preferences are indexed by the set of possible beliefs the decision-maker has regarding her opponent's move. Whilst this approach certainly addresses the issue concerning preferences over own mixed strategies, it also introduces new difficulties.

In particular, if preferences are to be represented by an expected utility function, they must satisfy a diversity condition. In [GS], this condition says that for every set of four alternatives available to the decision-maker, and each possible strict ordering of the four alternatives (there are $4! = 24$ of these), there exists a context such that preferences agree with that ranking. These diversity conditions are not only unnecessary for an expected utility representation, they are also strong enough to exclude the majority of possible decision-makers. Even the improvements we have made in chapter 2 exclude all possible preferences when $|S| = 2$ and $|A| \geqslant 3$, as well as very reasonable preferences that are not quite diverse enough.

Moreover, there appears to be no intuitive conditions on preferences that will resolve this issue unless, as in Chapter 3, we are willing to forego the integral representation and settle for an ordinal representation with connectedness and star-convexity replacing convexity. In this case, conditions such as diversity are superfluous. Whilst such representations are much more general and thereby capture a far wider variety of behavior, they face two drawbacks. The first is that the mathematical tools available for sums and integrals are sometimes indispensable for the application at hand. The second is that convexity gives us global properties of the utility function over context space, and this means we can elicit a utility function precisely with a relatively small amount of data.

It seems natural therefore to seek to identify a minimal extension of preferences that allows us to obtain an expected utility function without recourse to a diversity condition on preferences. I show that it suffices to define preferences over alternative-context pairs. The results are obtained in the setting where, for each alternative, the contexts are a mixture space. This latter concept was introduced in the classic paper of Herstein and Milnor (1953) [HM]. This added level of generality provides a common theoretical framework with which to model a variety of decision problems that are not special cases of the canonical example studied throughout this thesis. Two of these are discussed at the end of the chapter: state-dependent preferences and the model of Karni and Safra (2000) and the Allais paradox. For the latter, the set of mixture spaces need not be exogenously defined. Indeed, perhaps the main contribution of this chapter is to make progress toward removing this extra-behavioural condition that is present in all the models of which I am aware.

4.1.1 Two motivating examples

The first class of problems that is well suited to the model I present in this chapter is where the decision-maker's "opponent" is nature. Moreover,

(i) she knows she *will* face a choice in the presence of uncertainty about the future state of nature;

(ii) she knows, that when the situation arises, she will have knowledge of the context she is in (for instance, she will know the likelihood of any given state of nature);

(iii) given a context, she knows her preference for one alternative over another;

(iv) given a particular course of action, she knows which context she would rather be in;

(v) building upon (iii) and (iv), she is willing to go further and make statements of the form "I prefer to choose alternative a when in context p, than to choose alternative b when in context q".[1]

Although (ii) of this list is arguably a strong assumption for the class of examples of this chapter, I will simply take it as given and accept it as a topic for future research without further discussion.

Example 4.1. *Consider a planner who is devising a complete, contingent plan of how to respond to the future threat of a flood. There are two states (flood and no-flood) and there are two alternatives: do nothing or evacuate. Crucially, the contingencies are defined to be the set of possible probability distributions over states, not the states that may subsequently obtain. The idea being that, when the time comes, the action should be carried out without question.*

In this example, the planner should have no reason to consider mixtures over her set of alternatives: there is no obvious benefit to doing so. The situation is different from a game against a strategic opponent, such as rock-paper-scissors, where being predictable carries a cost. As Rubinstein (2000) and [GS] highlight, data on players' preferences over mixed strategies may be unreliable as it is not clear whether observable, pure actions of the player are part of a grand mixed strategy. If this hypothesis is true in games against other players, it is

[1] In this chapter we take this statement to be equivalent to "I prefer to be in context p when choosing alternative a, than to be in context q when choosing alternative b". Also, note that the above statement is a preference over verbs or actions each coupled with a context and seems both closer to our standard use of language and more accurate in its meaning than "I prefer alternative a in context p over alternative b in context q".

even more true in games that are “against nature”.

On the other hand, there is good reason for the planner to consider each possible contingency (context) and, given this, say whether she would evacuate or not. It is not too much more to ask her to state her preferences over contingencies given a choice of alternative. This suggests that (iii) and (iv) are reasonable assumptions in the flood example.

This leaves (v). Now the planner needs to be willing to make statements of the form “I prefer to announce evacuate when the probability of flooding is $\frac{1}{2}$ than to make no announcement when the probability of flooding is $\frac{1}{4}$”. Whilst there is no doubt that this is more demanding, it is worthwhile putting it into perspective by considering the following class of statements “I prefer to announce evacuate with probability $\frac{1}{2}$ when the probability of flooding is $\frac{1}{2}$ than to evacuate with probability $\frac{1}{4}$ when the probability of flooding is $\frac{1}{4}$”. The latter type of statement is necessary if we wish to apply the benchmark [vNM] model of expected utility.

In fact the [vNM] model requires quite a bit more than this. It requires that the decision-maker is willing to make preference statements about any pair of probability distributions over the set of four outcomes defined by taking the product of the set of alternatives with the product of the set of states: (d,n)≡(do nothing, no flood), (d,f)≡(do nothing, flood), (e,n)≡(evacuate, no flood) and (e,f)≡(evacuate, flood).

This space of lotteries contains probability distributions where the (joint) probability of the outcome (evacuate, flood) is not equal to the product of the marginal probabilities. Let δ_x denote the probability measure assigning probability one to outcome x and consider the following lottery:

$$\tfrac{1}{2}\,\delta_{(d,n)} + \tfrac{1}{4}\,\delta_{(d,f)} + \tfrac{1}{4}\,\delta_{(e,n)}.$$

For outcome (d,n) to occur with probability $\frac{1}{2}$, it seems reasonable to assume, that in the absence of mischievous deities, there is a positive probability that the central planner chooses "do nothing" *and* a positive probability that "flood" occurs. Similarly, for (e,n) to occur with probability $\frac{1}{4}$, the planner ought to be choosing "evacuate" with positive probability. However, by taking the product of the marginal probabilities of evacuate and flood, outcome (e,f) occurs with positive probability. Now this can't be because the sum of $\frac{1}{2}$, $\frac{1}{4}$ and $\frac{1}{4}$ is 1.

Even so, suppose the planner was willing to entertain the possibility of her strategies somehow being correlated with those of nature's; would she be willing to say whether she preferred the above lottery to

$$\tfrac{7}{16}\,\delta_{(d,n)} + \tfrac{3}{16}\,\delta_{(d,f)} + \tfrac{1}{16}\,\delta_{(e,n)} + \tfrac{5}{16}\,\delta_{(e,f)}\,?$$

It seems that a simple but challenging decision problem has been turned into a complicated problem that is even more challenging. The difficulty is that to define preferences over the space of lotteries on $A \times S$, where A is the set of alternatives and S is the set of states, is to define preferences over the 3-dimensional unit simplex $\Delta(A \times S)$ in $\mathcal{R}^4$. This adds considerable complexity to the decision problem.

One solution to this latter difficulty is to use the multilinear expected utility model of Fishburn (1980). This model allows us to define preferences on the product of unit-simplices $\Delta(A) \times \Delta(S) = [0,1] \times [0,1]$. It therefore allows the decision-maker to avoid defining her preferences on the strange lotteries described above. However, it still requires that the decision-maker define her preferences over her own mixed strategies. As a result, the conditions on preferences that Fishburn imposes are more numerous and more complicated than the model of this chapter.

Instead, I propose to define a simple model of preferences over the product $A \times \Delta(S)$. In the above example this amounts to two copies of the unit interval. Aside from the intuitive appeal in decision problems like the flood example above, the advantage to defining preferences on a smaller space, in terms of the complexity of eliciting a utility function may be significant in real-life complex decision problems such as those studied by computer scientists like Braziunas and Boutilier (2010).

In order to show that the present model is not restricted to choice under risk/ uncertainty, I now present a second example where the space of contexts need not be the set of probability distributions over states. It serves to motivate the concept of a mixture space which is defined formally below. For now it suffices to think of it as a suitable generalization of a convex space in which the operation of taking mixtures is defined.

Example 4.2. *Consider a customer in a restaurant choosing a meal with the assistance of a waiter. The menu defines a finite set of alternatives, and there are a continuum of possibilities for each dish: how well the flavors of the main ingredients combine, the degree to which the food will be cooked, the quantity of salt it will contain, etc. Suppose that the waiter is able to describe the context space as precisely as the customer wishes.*

Given the choice of a particular item on the menu, the customer may well be willing to state her preferences over the possible contexts, thus (iv) may be reasonable enough here. She may also be willing to make statements of the form: "I prefer the carrot soup with a table spoon and a half of cream, to the tomato salad with one quarter of an onion in it."

On the other hand, it is unclear that the decision-maker would be comfortable defining her preferences over mixtures of "soup" and "salad", even if by mixture we did mean probability mixtures. Presumably the only situation where she might even consider such alternatives is when she is unsure of what to do and flips a coin to break the tie. (So Fishburn's multilinear model, which also holds for mixture spaces, is arguably inappropriate here.) As above, to define preferences on a single mixture space, we must go even further: either assume the decision-maker considers lotteries over the product of the alternative space with the space of contexts; or allow for a continuum of portion sizes for all items, and expand the menu to include mixtures of all possible ingredients, cooking styles etc.

At this point it is natural to question the need for introducing any of the above strange considerations to the decision problem. The answer is that we are often interested in representing preferences with a utility function for which numerical values have meaning in same sense that temperature values do. That is given a choice of scale and origin, say, Celsius, we can speak of temperature as a number. Moreover, if someone else uses another scale and origin, such as Fahrenheit, their statements are meaningful to us because we know the affine transformation that converts one to the other.[2]

In game theory for instance the starting point is to specify a given player's "pay-offs" as numbers. [vNM] showed that this is a reasonable starting point provided preferences are defined on the set of probability distributions (or lotteries) over the outcomes of the game, and provided certain conditions on preferences apply. In this case, the player's payoffs are unique up to a common scale and origin (a positive affine transformation).

For the customer in the restaurant example, where lotteries may play no part, if we seek such a utility representation, we must appeal to the generalization of [vNM] by [HM]. They define preferences over a mixture space and impose very similar conditions on preferences to those of [vNM]. Their representation is also unique up to a positive affine transformation, and so utility units in their model are meaningful.

4.1.2 Synopsis and outline

The main purpose of the present chapter is to pin down the minimal conditions on preferences that extend [HM] (and hence [vNM]) to the setting where there

[2]The analogy is not so good if we consider Kelvin as this has a fixed origin.

is more than one mixture space and provide a precise characterization of the utility function.

The model of Karni and Safra (2000) (henceforth [KS]) comes closest to the present class of problems. There are two reasons why I choose to build a new model. The first is that they introduce additional structure on the space on which preferences are defined. This structure is not needed to make the extension of [vNM] and [HM] to the general setting where mixtures are not everywhere defined. I apply the model of the present chapter to the space they define in section (4.4.1). Second, they impose conditions on preferences that are have no counterpart in the original models of [vNM] and [HM], and are specific to the space on which they define preferences. This means that the representation in this chapter is therefore both simpler and more general than [KS] and the related paper of Karni (2009). Correspondingly, it is less straightforward to obtain, and the length of the derivation in section (4.3) testifies to this.[3]

What distinguishes the utility function of this chapter from that of [vNM], [HM], Fishburn (1980), [KS], but also the "state-dependent" utility models, is the extent to which utility is numerically comparable over the domain of preferences. (In "state-dependent" utility models (eg. Dréze (1961)), comparing utility numbers across states is meaningless, for there is an independent scale and origin to utility for each state. It is the polar opposite of [vNM] and the

[3]Moreover, the relevant result in that paper contains two mistakes. The first is their claim that the utility function is unique up to a common scale but not a common origin. As Karni (2009) points out and amends, it is in fact unique up to a positive affine transformation. The second mistake is also present in Karni (2009) and a counterexample to their claim is presented in section (4.4.1) below.

other models we have discussed.)

In the model of this chapter there is a spectrum of possibilities. For some decision-makers, utility is numerically comparable across mixture spaces in the sense that the utility function is, to use the terminology of Karni (2009), unique up to a positive affine transformation that applies uniformly across the domain (i.e. utility is cardinally measurable and fully comparable across the domain (CFC)). For others, the utility function will only be CFC within certain subsets of its domain. If such a subset is maximal, in the sense that it is the largest subset for which numerical utility comparisons are possible, we will call it a quasi-component of preferences.[4] Across quasi-components, only ordinal utility comparisons are meaningful.

Such concepts may at first sight seem irrelevant to decision-making. However the essential idea is that when, regardless of context, one alternative is obviously better than another, it may be that the decision-maker has *no* need for the high resolution measurement scale that [vNM] require. For other pairs of alternatives, the [vNM] model may be natural. The issue bears resemblance to the way that one does not typically need scales to decide whether the typical toddler is lighter than the typical adult, but periodically, we do need precise scales to measure whether our own weight has increased or decreased.

An application of such reasoning may help to explain the Allais Paradox. is provided in section (4.4.2), the final section of this chapter, where I will

[4] This terminology is related to the concept of a component in topology. These sets resemble components in some ways, but since they may contain limit points of other quasi-components, we have chosen this terminology.

argue that, when preferences are incompletely defined on the space of lotteries over monetary outcomes, preferences that accord to the Allais paradox are straightforwardly captured by a mixture-preserving utility function that is not defined over the entire simplex of lotteries. This application also highlights the fact that in general the set A need not index alternatives, it may index the members of a union of subsets of a single simplex.

The next section presents the conditions we impose upon preferences, the mixture space as well as the space of alternatives. Following this, in section (4.3) the two main representations are derived, the first generalizes the model of [HM], and the second the model of [vNM].[5]

The chapter then concludes with two applications of the model. The first, in section (4.4.1) makes precise the differences between our model and that of [KS] and shows how state-dependent preference can hold despite utility being CFC across states. I then conclude in section (4.4.2) with the application to the Allais Paradox.

4.2 Limitations of the model

To keep the notation as simple as possible, I assume that the space of contexts, $\mathcal{M}$, is the same for all the elements of A, which we will refer to simply as alternatives. Moreover $\mathcal{M}$ is a mixture space, which is defined in the following way:

[5]Although great care has been taken to define every new concept and explain each step in the proofs, the derivation is rather technical. Hence, I suggest that non-specialists first read the following section, which makes precise the limitations of the model, and continue reading near the end of section (4.3) where the representations are to be found.

Definition 4.1 (Mixture set). *A set $\mathcal{M}$ is said to be a* mixture set (or space) *if for any $x, y \in \mathcal{M}$ and any λ, we can associate another element, which we write as either $\lambda x + (1 - \lambda)y$ or $x\,\lambda\,y$, which is again in $\mathcal{M}$, and where*

(1) $x\lambda x = x$

(2) $x\lambda y = y(1 - \lambda)x$

(3) $(x\lambda y)\mu y = x(\lambda\mu)y$.

A mixture space is more general than a simplex. As Mongin (2001) shows, a mixture space needn't even be isomorphic to convex subset of a vector space. Even so, the following pair of examples shows that it is not general enough to include spaces such as $A \times \mathcal{M}$, where A is a general set of alternatives. The issue is of course that the mixture operation need not be defined on the whole of $A \times \mathcal{M}$.

Example. *Let $A = \{a, b\}$. Let (a, p) be an element of $\{a\} \times \mathcal{M}$ and $y = (b, p)$ an element of $\{b\} \times \mathcal{M}$. Clearly, there is no $\lambda \in [0, 1]$ other than 0 and 1 such that $x\lambda y$ lies in $A \times \mathcal{M}$.*

Example. *If A is the set of rational numbers the set $A \times \mathcal{M}$ is neither a mixture space, nor a product of mixture spaces. For if we are to take the interval $[0, 1]$, then $0 < \frac{1}{\sqrt{2}} < 1$ is not rational. Neither therefore is $(\frac{1}{\sqrt{2}}, p) \in A \times \mathcal{M}$ for any $p \in \mathcal{M}$.*

Note that the set we study can also be written as a union as follows

$$A \times \mathcal{M} \equiv \bigcup_{a \in A} (\{a\} \times \mathcal{M}).$$

Unlike [HM], I will assume that the mixture space $\mathcal{M}$ is endowed with a topology, and that, under this topology, it is compact. This, together with the conditions I impose on preferences, as I show below, is sufficient for the existence of a greatest lower bound (glb) and least upper bound (lub) of $\mathcal{M}$ under the order that preferences define for each alternative $b \in A$.

This is an important simplification that makes the proofs more straightforward. Further research is required to prove that this condition can be weakened so as to have a representation for a union of general mixture spaces.[6] Nonetheless, it still allows for a wide range of context spaces as the following example highlights.

Example 4.3. *For an arbitrary set S, $[0,1]^S$ is, by the Tychonov theorem, compact in the product topology. Now since $\Delta(S)$, is a closed subset of $[0,1]^S$, by theorem 26.2 of Munkres p.181 $\Delta(S)$ is also compact, and since it is convex, it is a mixture space and thus a valid space of contexts for the results that follow.*

Unless otherwise stated, the space of alternatives will be assumed to be finite. Where possible my proofs are written so that they would also apply to the countably infinite case. Preliminary research into the latter shows that it is somewhat more complicated, and that to progress we will need to make slight alterations to some of the concepts we introduce. For the case where A is uncountable, it is clear that further conditions on preferences will be needed. The reason being that there can only be countably many disjoint intervals (with nonempty interior) in $\mathcal{R}$, whereas in general preferences may lexicographically order elements of $A \times \mathcal{M}$ in the sense that for each $a, b \in A$, either (a, p) is

[6] Although these results are not presented here, my current efforts appear to show that, for a topological mixture space at least, this should be possible.

strictly better than (b,q) for all $p,q \in \mathcal{M}$, or the reverse strict preference for all $p,q \in \mathcal{M}$, so that any representation of such preferences will need to map into the same number of disjoint intervals as the cardinality of A.

For any given $x,y \in A \times \mathcal{M}$, we take the statement "y is weakly preferred to x" to be equivalent to $x \precsim y$.[7] The conditions on the relation $\precsim$ that will be needed are defined as follows.

Definition (Complete pre-order (O))**.**

For all $x,y,z \in A \times \mathcal{M}$, both the following hold:

*(i) (*Completeness*) $x \precsim y$ or $y \precsim x$, and*

*(ii) (*Transitivity*) if $x \precsim y$ and $y \precsim z$, then $x \precsim z$.*

Definition (Continuity (C'ty))**.**

For all $x,y \in A \times \mathcal{M}$, the following sets are closed:

$$\{y \in A \times \mathcal{M} : y \precsim x\}, \quad \textit{and} \quad \{y \in A \times \mathcal{M} : x \precsim y\}$$

When the decision-maker makes a preference statement comparing different contexts $p,q \in \mathcal{M}$, *given that she is choosing a particular alternative $b \in A$*, I will use the shorthand $p \precsim_b q$ and understand it to be the same statement as $(b,p) \precsim (b,q)$.

Definition ([vNM)**.** *Independence on each mixture space (I)]*

For any $b \in A$, and any $p,q,r \in \mathcal{M}$ if $p \sim_b q$, then

$$p\tfrac{1}{2}r \sim_b q\tfrac{1}{2}r.$$

[7]This notation is used by Fishburn (1979) and Binmore (2009). I believe it to be easier to read in the present setting due to the fact that our proof often deals with "intervals" defined by preferences.

Condition (I) is stated in the form that [HM] introduced in their paper. The first part of that paper is dedicated to showing that this implies the more familiar form that [vNM] introduced, where the condition in (I) holds not only for $\lambda = \frac{1}{2}$, but for all $\lambda \in [0, 1]$. The following condition is to my knowledge new, and it is the condition that as I now show generalizes condition (I).

Definition (Congruent betweenness (CB))**.**
For any $b, c \in A$, $p, q, p', q' \in \mathcal{M}$, if both $(b, p) \sim (c, q)$ and $(b, p') \sim (c, q')$ then

$$(b, p\tfrac{1}{2}p') \sim (c, q\tfrac{1}{2}q').$$

Clearly if the cardinality of A is one, then condition (CB) is implied by condition (I) and transitivity; since (I) implies that whenever $p \sim_b q$ and $p' \sim_b q'$ we have

$$p\tfrac{1}{2}p' \sim_b q\tfrac{1}{2}p' \sim_b q\tfrac{1}{2}q',$$

and transitivity ensures that that indecisiveness propagates. In fact, by taking $p = q$ and $b = c$ in the definition of (CB), we see that, because $p = q$ implies $p \sim_b q$, (CB) implies (I). Thus, when the cardinality of A is one, the two conditions are equivalent in the presence of (O), but when A contains two or more elements, (CB) implies (I), but not vice versa.

This justifies my claim that (CB) is a natural generalization of [vNM] independence to unions of mixture spaces, or equivalently, spaces for which the mixture operation is not everywhere defined. Note that if for two mixture spaces each element of the first is strictly better than all elements of the second, then condition (CB) is silent for such comparisons, but it still has implications within each mixture space.

Note that [KS] provide an example that shows why their version of the (CB) is not implied by the combination of (O), (C'ty) and (I). The intuition is that whenever there are at least two distinct indifference sets, both of which contain elements from two particular mixture spaces, it is possible to find utility functions that are mixture preserving on each of the mixture spaces, together satisfy (O), (C'ty) and (I), but which fail to satisfy (CB).

If $|A| = 1$, then $A \times \mathcal{M}$ is in fact a mixture space, and if it satisfies conditions (O), (C'ty) and (CB), then we say that it is a vNM ordered space. More generally, we have the following definition.

Definition 4.2. *[Extended vNM ordered space]*
Let $\mathcal{M}$ be a mixture space and A a discrete set, and let $\precsim$ be a binary relation on $A \times \mathcal{M}$. Then $(A \times \mathcal{M}, \precsim)$ will be referred to as an extended vNM ordered space *if it satisfies conditions (O), (C'ty), and (CB).*

Any representation of preferences over a mixture set will involve a function that is, first and foremost, mixture preserving.

Definition 4.3 (A generalization of [HM] and Moulin (2001)[8])**.** *For any set D, a function $f : D \to \mathcal{R}$ is said to be* mixture preserving *(MP) if for all $p, q \in D$, and all $\lambda \in [0, 1]$ such that $p\lambda q \in D$*

$$f(p\lambda q) = \lambda f(p) + (1 - \lambda) f(q).$$

[8] I thank Peter Hammond for recommending this form of the definition.

A special case of a mixture preserving function is of course an expected utility function such as that of [vNM].

4.3 Mixture preserving utility

The next lemma ensures that for each $a \in A$, the vNM ordered space $\{a\} \times \mathcal{M}$, has a mixture preserving representation.

Lemma 4.1. *Let* $|A| = 1$*. Then* $(A \times \mathcal{M}, \precsim) \equiv (\{a\} \times \mathcal{M}, \precsim)$ *is a vNM ordered space if and only if there exists a mixture preserving function* $U : \{a\} \times \mathcal{M} \to \mathcal{R}$ *such that for every* $p, q \in \mathcal{M}$*,*

$$(a, p) \precsim (a, q) \quad \Leftrightarrow \quad U(a, p) \leqslant U(a, q) \tag{4.1}$$

Proof. The proof follows from the fact that when $A \times \mathcal{M}$ is a vNM ordered space it satisfies the conditions of [HM]. The only part that is not immediate, is the observation that condition (C'ty) is stronger than the continuity condition of [HM]. Their condition is stated as follows.
For any $p, q, r \in \Delta$, the following sets are closed:

$$\{\lambda \in [0, 1] : p\lambda q \precsim_a r\} \quad \text{and} \quad \{\lambda \in [0, 1] : r \precsim_a p\lambda r\}.$$

We prove the contrapositive. That is, we prove that if [HM]'s condition fails to hold, then so does (C'ty). Suppose there exists a sequence $\{\lambda_n : n \in \mathbb{N}\}$ in $[0, 1]$ that converges to λ' with the property that, for all n, $p\lambda_n q \precsim_a r$, whilst at λ' we have $r \prec_a p\lambda' q$. In the presence of complete preferences, this is the only possibility. Now, this implies that (C'ty) indeed fails to hold. □

Lemma (4.1) also provides justification for the following identity

$$(\{a\} \times \mathcal{M}, \precsim) \equiv (\mathcal{M}, \precsim_a)$$

and we will use the latter as shorthand for the former, thus $(a, p) \sim (a, q)$ if and only if $p \sim_a q$ for example.

The first result that arises from the assumption of compactness is the following.

Lemma 4.2. *Let $[\alpha, \gamma]$ be a closed interval in $\mathcal{R}$ such that $\alpha < \gamma$ and let $(\mathcal{M}, \precsim)$ be a vNM ordered space with $\prec \neq \varnothing$. Then given any mixture preserving representation $\widetilde{U}$ of $(\mathcal{M}, \precsim)$, there exists unique $\theta, \kappa \in \mathcal{R}$ with $\theta > 0$ such that the function*

$$U : \mathcal{C} \to [\alpha, \gamma], \quad p \mapsto \theta\widetilde{U}(p) + \kappa,$$

satisfies

$$p \precsim . \, q \quad \Leftrightarrow \quad U(p) \leqslant U(q)$$

for each $p, q \in \mathcal{M}$.

Proof of lemma (4.2). Since $(\mathcal{M}, \precsim)$ is a vNM ordered space, by lemma (4.1) there exists a mixture preserving function $\widetilde{U} : \mathcal{M} \to \mathcal{R}$ that represents preferences on $\mathcal{M}$. The existence of mixture preserving representation is sufficient for condition (C'ty), and so $\widetilde{U}$ is continuous. Then since $\mathcal{M}$ is compact, by theorem 26.4 of Munkres p182, the image of $\mathcal{M}$ under $\widetilde{U}$ is compact.

By the extreme value theorem (theorem 27.4 of Munkres 190), this implies that there exists a greatest lower bound g and a least upper bound l, both in

$\mathcal{M}$ such that

$$\widetilde{U}(g) \leqslant \widetilde{U}(p) \leqslant \widetilde{U}(l)$$

for every $p \in \mathcal{M}$. Let $\widetilde{\alpha} = \widetilde{U}(g)$ and $\widetilde{\gamma} = \widetilde{U}(l)$. By the fact that $\widetilde{U}$ is a representation, we know that for all $p \in \mathcal{M}$, $g \precsim p \precsim l$.

By condition (2) of the definition of mixture spaces, for every $\lambda \in [0,1]$ $g\lambda l$ is an element of $\mathcal{M}$. Then since $\widetilde{U}$ is mixture preserving, for all $\lambda \in [0,1]$

$$\widetilde{U}(g\lambda l) = \lambda\widetilde{\alpha} + (1-\lambda)\widetilde{\gamma},$$

so that the image of $\widetilde{U}$ is equal to the interval $[\widetilde{\alpha}, \widetilde{\gamma}]$. The fact that $\prec$ is nonempty, and implies that $l \prec g$ and hence $\widetilde{\alpha} < \widetilde{\gamma}$.

Now let θ satisfy the equation $\theta(\widetilde{\gamma} - \widetilde{\alpha}) = \gamma - \alpha$. Then $\theta > 0$ and it is uniquely identified. Next, let κ satisfy $\theta\widetilde{\alpha} + \kappa = \alpha$; it too is uniquely identified. Then since

$$\theta\widetilde{\gamma} + \kappa = \theta\widetilde{\gamma} + \alpha - \theta\widetilde{\alpha} = \gamma,$$

we see that $U := \theta\widetilde{U} + \kappa$ is a candidate for the required function.

The fact that U is mixture preserving follows readily from the fact that $\widetilde{U}$ is mixture preserving and the fact that for each $p, q \in \mathcal{M}$, $0 \leqslant \lambda \leqslant 1$

$$\begin{aligned} U(p\lambda q) :=& \theta\,\widetilde{U}(p\lambda q) + \kappa \\ =& \lambda\,(\theta\,\widetilde{U}(p) + \kappa) + (1-\lambda)(\theta\,\widetilde{U}(q) + \kappa) \\ =& \lambda\,U(p) + (1-\lambda)\,U(q); \end{aligned}$$

whilst the fact that it is order preserving is similarly easy to show. $\square$

Remark 4.1. *If $\prec$ is empty, then for all $p, q \in \mathcal{M}$, $p \sim q$ and so $\widetilde{U}(p) = \widetilde{\alpha}$ for all $p \in \mathcal{M}$. Clearly if α is any other element of $\mathcal{R}$, then there exists an infinity of solutions to the equation $\theta\widetilde{\alpha} + \kappa = \alpha$.*

Lemma 4.3. *Suppose that there exist $a \neq b$ in A and $p, q, p' \in \mathcal{M}$ such that preferences satisfy $(a,p) \precsim (b,q) \precsim (a,p')$ and $(a,p) \prec (a,p')$. Then there exists a unique $\lambda \in [0,1]$ such that $(a, p\lambda p') \sim (b,q)$.*

Proof. Existence: This argument parallels the proof of Theorem 1 of [HM]. Consider the the set

$$T := \{\lambda \in [0,1] : (b,q) \precsim (a, p\lambda p')\}.$$

By condition (C'ty) and the proof of lemma (4.1), T is a closed subset of $[0,1]$. Since $(b,q) \precsim (a,p')$, $1 \in T$; so T is nonempty. By the same argument, $W := \{\lambda \in [0,1] : (a, p\lambda p') \precsim q\}$ is also closed in $[0,1]$; it is nonempty as $0 \in W$. Now since $\mathcal{M}$ is a completely preordered mixture set, $T \cup W = [0,1]$; the fact that $[0,1]$ is a connected set, implies that no pair of its subsets define a separation thereof, and so $T \cap W$ is nonempty. Let $\lambda_0 \in T \cap W$; by construction of these sets I have $(a, p\lambda_0 p') \precsim (b,q) \precsim (a, p\lambda_0 p')$. So that by asymmetry of $\prec$ I have $(a, p\lambda_0 p') \sim (b,q)$.

Uniqueness: Clearly if $(a,p) \prec (b,q) \prec (a,p')$, then $0 < \lambda_0 < 1$. Now take any $r \in \mathcal{M}$ such that $(a,r) \sim (b,q)$. Transitivity of $\sim$ implies that $(a,r) \sim (a, p\lambda_0 p')$. Moreover, theorem 6 of [HM] together with the fact that I have assumed $(a,p) \prec (a,p')$ implies that λ_0 is unique. $\square$

Definition 4.4. *Let x and y be elements of $A \times \mathcal{M}$. Then the pair (x, y) is called a* gap *if both the following conditions hold:*

(i) $x \prec y$; and

(ii) the set $]x, y[:= \{z : x \prec z \prec y\}$ is empty.

Remark 4.2. *Note that more generally the set $]x, y[:= \{z : x \prec z \prec y\}$ can, by de Morgan's laws and the fact that $\precsim$ is complete and transitive, be rewritten as $\{z : \neg(z \precsim x)\} \cup \{z : \neg(y \precsim z)\}$. Then condition (C'ty) implies that each of the sets in this union is open; therefore $]x, y[$ is open.*

Definition 4.5. *Let $x = (a, p)$ and $y = (b, q)$ be elements of $A \times \mathcal{M}$. Then the pair (x, y) is called a* quasi-gap *(or alternatively a* q-gap*) if it is either a gap or if both the following conditions hold:*

(i) $x \sim y$; and

(ii) the sets

$$D_x = \{c \in A : \text{for some } r \in \mathcal{M}, (c, r) \prec x\}; \text{ and}$$
$$E_y = \{c \in A : \text{for some } r \in \mathcal{M}, y \prec (c, r)\}$$

are disjoint with $a \in D_x$ and $b \in E_y$.

Definition 4.6. *In what follows I will denote any element that is a greatest lower bound for a set $\{c\} \times \mathcal{M}$ by g_c. Similarly, l_c will refer the the least upper bound.*

Lemma 4.4. *If (x, y) is a q-gap, then for all $c \in A$, either $(c, r) \precsim x$ for all $r \in \mathcal{M}$, or $y \precsim (c, r)$ for all $r \in \mathcal{M}$.*

Proof. If $x \sim y$, so that (x, y) is a q-gap but not a gap, then in property (ii) of the definition of a q-gap, the fact that the sets D_x and E_y are disjoint is sufficient for the conclusion of this lemma to hold.

If $x \prec y$, then the conclusion of this lemma is equivalent to the following:

$$\forall c \in A \quad \neg(x \prec l_c \text{ and } g_c \prec y)$$

which in turn, by the fact that $]x, y[= \varnothing$, is equivalent to:

$$\forall c \in A \quad \neg(g_c \precsim x \prec y \precsim l_c).$$

So by way of contradiction, suppose not. If $g_c \sim x \prec y \sim l_c$, then the fact that $\{c\} \times \mathcal{M}$ is a vNM ordered set together with theorem 2b of [HM] implies that there exists $z \in \{c\} \times \mathcal{M}$ such that $g_c \prec z \prec l_c$. Transitivity then implies that $x \prec z \prec y$, so that the pair (x, y) cannot be a gap; therefore, either $g_c \prec x$ or $y \prec l_c$ must hold.

Suppose that $g_c \prec x$, then the fact that $g_c \prec x \prec l_c$ together with by lemma (4.3) there exists $v \in \{c\} \times \mathcal{M}$ such that $v \sim x$. Repeating the argument of the previous paragraph leads us to the same contradiction. The case where $y \prec l_c$ is identical. $\square$

Lemma 4.5. *Every q-gap (x, y) of $(A \times \mathcal{M}, \precsim)$ has the property that $x \sim l_a$ and $y \sim g_b$ for some $a \neq b$ in A.*

Proof. Suppose that (x, y) is a gap, and suppose that $x \nsim l_c$ for all $c \in A$. Then there exists $d \in A$, $p, q \in \mathcal{M}$ such that $x = (d, p) \prec l_d = (d, q)$. Now since

$(\mathcal{M}, \precsim_d)$ is a vNM ordered space, by theorem 2b of [HM] I have $p \prec_d p\frac{1}{2}q \prec_d q$, so that because $]x, y[= \varnothing$, and $\precsim_d$ is complete, $y \precsim \big(d, p\frac{1}{2}q\big)$. Now the same theorem implies that $y \precsim \big(d, p\frac{1}{2}(p\frac{1}{2}q)\big)$ which, by conditions (2) and (3) of the definition of a mixture space

$$p\tfrac{1}{2}(p\tfrac{1}{2}q) = p\tfrac{1}{2}(q\tfrac{1}{2}p) = (q\tfrac{1}{2}p)\tfrac{1}{2}p = q\tfrac{1}{4}p = p(1 - \tfrac{1}{4})q.$$

Thus $y \precsim \big(d, p(1 - \frac{1}{4})q\big)$. Indeed by the induction hypothesis, and an identical argument I see that for all $j \in \mathbb{N}$ $y \precsim \big(d, p(1 - 2^{-j})q\big)$, so that by condition (C'ty) and the fact that $(d, p(1 - 2^{-j})q)$ converges to $x = (d, p)$, we see that $x \sim y$. This implies that (x, y) cannot be a gap; this contradiction implies that there exists $a \in A$ such that $x \sim l_a$. An identical argument shows that $y \sim g_b$ for some $b \in A$ whenever (x, y) is a gap. Moreover, by the definition of a gap, $a \neq b$.

Now suppose that $x \sim y$, that is (x, y) is a q-gap, but not a gap. By property (ii) of q-gaps, there exists $d \in E_y \backslash D_x$ with $y = (d, p)$ for some $p \in \mathcal{M}$. The fact that d lies outside D_x implies, by condition (O), that $x \precsim (d, r)$ for all $r \in \mathcal{M}$, so that $y \sim g_d$. Finally, by the same property of q-gaps, there exists $(c', p') \in A \times \mathcal{M}$ such that $(c', p') = x$ and $c' \in D_x \backslash E_y$. Now the fact that c lies outside E_y implies that for all $r \in \mathcal{M}$, $(c, r) \precsim y$, so that by $(c, r) \precsim (c, p')$ for all $r \in \mathcal{M}$. This implies that $x \sim l_c$; moreover, I can see that $d \neq c$. $\square$

I will now use the quasi-gaps of $(A \times \mathcal{M}, \precsim)$ to define quasi-components. These are subsets of $A \times \mathcal{M}$ upon which my eventual representation will have a common scale and origin. That is, on a given quasi-component, in the language of measurability and comparability, preferences are cardinally measurable and fully comparable.

Definition 4.7 (Quasi-component)**.** *Let (x, y) be a quasi-gap in $(A \times \mathcal{M}, \precsim)$.*

(i) If for all other q-gaps (w, z), $x \precsim w$, then the (nonempty) set

$$[\leftarrow, x] := \{t \in A \times \mathcal{M} : x \precsim t\}$$

is called a quasi-component or q-component of $(A \times \mathcal{M}, \precsim)$.

(ii) If for all other q-gaps (w, z), $z \precsim y$, then the (nonempty) set

$$[y, \rightarrow] := \{t \in A \times \mathcal{M} : y \precsim (a, p)\}$$

is called a quasi-component of $(A \times \mathcal{M}, \precsim)$.

(iii) If (u, v) is another q-gap, distinct from (x, y) with $y \prec v$, and for all other q-gaps (w, z), $z \precsim y$ or $u \precsim w$, then the (nonempty) set

$$[y, u] := \{t \in A \times \mathcal{M} : y \precsim t \precsim u\}$$

is called a quasi-component or q-component of $(A \times \mathcal{M}, \precsim)$.

(iv) If there are no q-gaps in $(A \times \mathcal{M}, \precsim)$, then the set $A \times \mathcal{M}$ is itself a quasi-component.

Definition 4.8 (Component)**.** *If the quasi-gap(s) that identify a quasi-component $[u, v]$ are gap(s), then $[u, v]$ is also called a component.*

Definition 4.9. *A quasi-gap (w, z) is said to be distinct from another quasi-gap (x, y) if either $w \prec x$ or $y \prec z$. One quasi-component is said to be distinct* *from another if at least one of the quasi-gaps that define it is distinct from each of the quasi-gaps of the other. A collection of distinct quasi-components is such that every pair in the collection is mutually distinct.*

The following lemma together with the fact that for every q-gap (x, y) the set $]x, y[$ is empty shows that the definition of q-components is such that for all $x \in A \times \mathcal{M}$, x belongs to some q-component.

Lemma 4.6. *For every q-gap (x, y) in $(A \times \mathcal{M}, \lesssim)$, there exists $w, z \in A \times \mathcal{M}$ such that $[w, x]$ and $[y, z]$ are q-components.*

Proof. If for every q-gap (w', w) in $(A \times \mathcal{M}, \lesssim)$, $x \lesssim w'$, then by (i) of the definition of q-components, $[\leftarrow, x]$ is a q-component. The fact that $[\leftarrow, x] = [w, x]$ for some $w \in A \times \mathcal{M}$ follows from the fact that A is finite and $\mathcal{M}$ is compact. The same is true of $[y, \rightarrow]$ if $w \lesssim y$ for every q-gap (w', w).

Now suppose that there exists a q-gap (v', v) with $v' \prec x$. I will first show that there exists a q-gap (w', w) with the property that every q-gap (u', u) with $v' \lesssim u' \lesssim x$ satisfies $u' \lesssim w'$.

Suppose not. That is, for every q-gap (w', w) such that $w' \prec x$, there exists another (u', u) with $w' \prec u' \prec x$. Starting with v', I may, using the first element in the pair that defines each of these q-gaps, construct a sequence $\{v'_i : i \in \mathbb{N}\}$ with the property that $v'_1 = v'$ and for all $i \geqslant 2$, $v'_{i-1} \prec v'_i \prec x$. By assumption this sequence is infinite. However, by lemma (4.5) I know that for all i, $v'_i \sim l_{a(i)}$ for some $a(i) \in A$; now the fact that $v'_{i-1} \prec v'_i$ for each i implies that A is an infinite set, a contradiction of my assumption that it is finite.

Thus there exists a q-gap (w', w) such that every other q-gap lies outside the set $[w, x]$; part (iii) of definition (4.7) then states that this is a q-component. $\square$

Lemma 4.7. *For all $a \in A$, $\{a\} \times \mathcal{M}$ is a subset of some q-component $[u, v]$. Moreover, if $\prec_a \neq \emptyset$, then $\{a\} \times \mathcal{M}$ belongs to at most one (distinct) q-component, otherwise it belongs to at most two.*

Proof. Lemma (4.6) implies that $x = (a, p) \in A \times \mathcal{M}$ lies in some q-component $[u, v]$. In turn, lemma (4.4) implies that the q-gaps that define $[u, v]$ must lie outside $\{a\} \times \mathcal{M}$, and this is sufficient for the first part of lemma.

It is also sufficient for the the fact that $\{a\} \times \mathcal{M}$ belongs to $[u, v]$ alone whenever $\prec_a$ is nonempty. For in this case, even if for instance $g_a \sim u$, there exists $q \in \mathcal{M}$ such that $g_a \prec (a, q)$, so that (a, q) cannot belong to any q-component in the set $[\leftarrow, u]$. In the same way, we know that $l_a \precsim v$, so that $(a, q) \notin [v, \rightarrow]$.

To see that when $\prec_a = \emptyset$, the set $\{a\} \times \mathcal{M}$ can belong to more than one q-component, note that if $g_a \sim u$ and $[y, u]$ is also a q-component for some $y \in A \times \mathcal{M}$, then the fact that $l_a \sim u$ implies that $\{a\} \times \mathcal{M}$ belongs to both q-components. For any other q-component $[w, z]$ that is distinct from both $[y, u]$ and $[u, v]$, either $w \prec y$ or $v \prec z$, in either of these cases, one cannot have $u \in [w, z]$. $\square$

Remark 4.3. *By this last lemma, denote any q-component $[u, v]$ as $B \times \mathcal{M}$ for some $B \subset A$ such that there are no q-gaps in $B \times \mathcal{M}$. Moreover, also by this lemma, for some $Q \in \mathbb{N}$ I may write*

$$A \times \mathcal{M} = \bigcup_{i=1}^{Q} B_i \times \mathcal{M},$$

where for each i, $B_i \times \mathcal{M}$ is a quasi-component.

Lemma 4.8. *For any q-component $[u, v] := B \times \mathcal{M}$, if $u \prec v$, then there exists $b \in B$ such that $\prec_b \neq \emptyset$ and $u \sim g_b$. If on the other hand $u \sim v$, then $[u, v]$ is a component.*

Proof. Suppose that despite the fact that $u \prec v$, it holds that for all $b \in B$, the set $\prec_b$ is empty. By this assumption, together with lemma (4.5), we know that $u \sim l_a \sim g_a$ and $v \sim g_c \sim l_c$ for some $a, c \in A$. Now since $[u, v]$ contains no q-gaps, there exists $d \in B \backslash \{a, c\}$ such that $g_a \prec g_d \sim l_d \prec g_c$. In turn, there exists $e \in B \backslash \{a, c, d\}$ such that $g_a \prec g_e \sim l_e \prec g_d$. A finite iteration of this argument exhausts the elements of the finite set B, so that we contradict the fact that there are no gaps in $[u, v]$; I conclude that there exists $b \in B$ with $\prec_b \neq \emptyset$.

An identical argument to the preceding paragraph shows that it is not the case that for all $b \in B$ with $\prec_b \neq \emptyset$ we have $u \prec g_b$; thus $u \sim g_b$ for some such b.

Suppose that $u \sim v$ and $[u, v]$ is a q-component but not a component, that is the q-gaps that identify $[u, v]$ are not gaps. Then let (x, u) (v, y) be quasi-gaps with $x \sim u$ and $v \sim y$. Transitivity, via condition (O), then implies that these two quasi-gaps are not distinct and so $[u, v]$ is not a q-component. □

Lemma 4.9. *The number Q of (distinct) quasi-components in $(A \times \mathcal{M}, \precsim)$ is less than or equal to the cardinality of A. The number of (distinct) quasi-gaps is one less than the number of quasi-components.*

Proof. By lemma (4.7) we know that every q-component $[u, v] = B \times \mathcal{M}$ contains at least one set $\{b\} \times \mathcal{M}$. If if $u \prec v$ then by lemma (4.8) $g_b \prec l_b$ for some

$b \in B$ and by lemma (4.7) therefore, $[u, v]$ is the only q-component to which the set belongs. If $u \sim v$, then $g_b \sim l_b$ for all $b \in B$, and $[u, v,]$ is a component, and in this case every b in the nonempty set B satisfies the property that $[u, v]$ is the only q-component to which $\{b\} \times \mathcal{M}$ belongs.

This in itself is sufficient for the proof of the first part of this lemma. For the second part, I proceed by induction. First consider the "lowest" q-component in the order $\precsim$ which is denoted by $[\leftarrow, v_1]$. Either there are no q-gaps in $(A \times \mathcal{M}, \precsim)$, so that $A \times \mathcal{M}$ is the only q-component, or there exists y_1 such that (v_1, y_1) is a q-gap. In the former case the statement of this lemma is true, and in the latter there are once more two possibilities. By lemma (4.6), there exists v_2 such that $[y_1, v_2]$ is a q-component, either $[y_1, v_2] = [y_1, \rightarrow]$ is a q-component, or there exists y_2 such that (v_2, y_2) is a q-gap. Once again, in the former case, the statement of the lemma is true, and in the latter there are two similar possibilities. It is clear that for each $2 \leqslant j \leqslant |A|$ the inductive step is identical to the above and is therefore omitted. $\square$

Lemma 4.10. *If* $[u, v] = B \times \mathcal{M}$ *is a q-component, then either* $u \sim v$ *or there exists a minimal sequence* $\{a_1, \dots a_h\}$ *in* A *such that for all* $j = 1, \dots, h$, $\{a_j\} \times \mathcal{M}$ *is a subset of* $[u, v]$*, and*

$$g_1 \prec g_2 \prec l_1 \precsim g_3 \prec l_2 \precsim g_4 \prec \cdots \prec g_{h-1} \prec l_{h-2} \precsim g_h \prec l_{h-1} \prec l_h.$$

Definition 4.10. *The set* $\{a_1, \dots, a_h\} \times \mathcal{M}$ *is called a* strict cover *of* $[u, v]$.

Proof. Let $u \prec v$, then by lemma (4.8) we may suppose that $g_b \prec l_b$ for some $b \in B$ and take g_b to be the glb of $[u, v]$. By the fact that B is finite together with condition (O), choose b such that for every c with $g_c \sim u$, $l_c \precsim l_b$. Consider the set

$$L_b := \{a \in A : g_a \prec l_b \prec l_a\}.$$

First observe that $L_b \subset B$, for I know that $l_b \precsim v$. If L_b is empty, then for every $a \in A \backslash L_b$, either $l_b \precsim g_a$ or $l_a \precsim l_b$. The former of these two relationships implies that if for some $p \in \mathcal{M}$, $l_b \prec (a, p)$ then $l_b \precsim g_a$, whilst together they imply for all $a \in B$, $l_a \precsim l_b$. Thus whenever L_b is empty $[u, v] = [g_b, l_b]$ and the sequence I seek is simply $\{g_b, l_b\}$ with $h = 1$.

If L_b is nonempty, then let $a_1 := b$, $g_1 := g_b$ and $L_1 := L_b$. Then take $a_2 \in L_b$ and $l_2 := l_{a_2}$ to satisfy $l_c \precsim l_2$ for all $c \in L_1$. Such an element exists because of condition (O) and the fact that L_1 is nonempty and finite. Now consider the set

$$L_2 := \{a \in A : g_a \prec l_2 \prec l_a\}.$$

By the same argument as for L_1 above, I see that $L_2 \subset B \backslash \{a_1\}$. If it is empty I have found a strict cover:

$$u \sim g_1 \prec g_2 \prec l_1 \prec l_2.$$

If not, then let a_3 be the element of L_2 satisfying $l_a \precsim l_3$ for all $a \in L_2$. If L_3, defined recursively as for L_1 and L_2 is empty then I claim my sequence is

$$g_1 \prec g_2 \prec l_1 \precsim g_3 \prec l_2 \prec l_3.$$

The only relationship that needs further explanation is $l_1 \precsim g_3$. This holds because otherwise $a_3 \in L_1$ and the fact that $l_2 \prec l_3$ would contradict the fact that l_2 was maximal. The general case follows by induction, the argument of which is identical to the one just given and is hence omitted.

The fact that the strict cover is minimal follows immediately from the construction. $\square$

Lemma 4.11. *If* $(a,p) \sim (b,q)$ *and* $(a,p') \sim (b,q')$*, then, for all* $0 < \lambda < 1$*,* $(a, p\lambda p') \sim (b, q\lambda q')$.

Proof. Without loss of generality, let $(a,p) \prec (a,p')$, so that, by transitivity, $(b,q) \prec (b,q')$. By condition (CB), $(a, p\frac{1}{2}p') \sim (b, q\frac{1}{2}q')$. Successive applications of this condition show that for all dyadic rational numbers, $0 < \pi = \sum_{i=1}^{n(\pi)} \zeta_i/2^i < 1$, where $\zeta_i = 0$ or 1, we have $(a, p\pi p') \sim (b, q\pi q')$. For the remainder of this proof, π will refer to the binary expansion of some dyadic rational number. Recall the fact that the set of such numbers is dense in the real numbers and so every $0 < \lambda < 1$ is the limit of some such sequence $\{\pi_j : j \in \mathbb{N}\}$.

For $0 < \lambda < 1$ consider the set $T_b := \{x \in A \times \mathcal{M} : x \precsim (b, q\lambda q')\}$. If $\{\pi_j\}$ is such that $\pi_j < \lambda$ for each j and $\lim_j \pi_j = \lambda$, then $(b, q\pi_j q') \prec (b, q\lambda q')$ for all j by theorem 4 of [HM]. By condition (BC) and transitivity, $(a, p\pi_j p') \prec (b, q\lambda q')$ for each j and so $\{(a, p\pi_j p')\} \subset T_b$. By condition (C'ty) T_b is closed, and because $\lim_j p\pi_j p' = p\lambda p'$ this implies that $(a, p\lambda p') \precsim (b, q\lambda q')$.

By the same argument $T_a := \{x \in A \times \mathcal{M} : x \precsim (a, p\lambda p')\}$ is closed and contains the set $\{(b, q\pi_j q')\}$ which converges to $(b, q\lambda q')$. Thus $(b, q\lambda q') \precsim (a, p\lambda p')$, so that by asymmetry of $\prec$ the proof is complete. $\square$

Theorem 4.1 (Mixture preserving representation on q-components). *$B \times \mathcal{M}$ is a quasi-component of the* extended *vNM ordered space $(A \times \mathcal{M}, \precsim)$ if and only if both the following conditions hold.*

(i) There exists a mixture preserving function $U : B \times \mathcal{M} \to \mathcal{R}$ such that for all $x, y \in B \times \mathcal{M}$

$$x \precsim y \quad \Leftrightarrow \quad U(x) \leqslant U(y).$$

(ii) If V is any other function with the same properties as U, then, for some $\theta, \kappa \in \mathcal{R}$ with $\theta > 0$, $V = \theta U + \kappa$.

Proof. If for all $x, y \in B \times \mathcal{M}$, $x \sim y$, then take U to satisfy $U(\cdot) \equiv 1$ and remark (4.1) ensures that every other representation V is the form $V = \theta U + \kappa$ for a one-dimensional set of suitable θ, κ combinations. In the opposite direction, suppose that $B \times \mathcal{M}$ is not a quasi-component. Then by lemma (4.7) $B \times \mathcal{M}$ is a union of q-components and so it contains at least one q-gap which, by lemma (4.5), which I denote by (l_a, g_b) for some $a, b \in B$. Now by the definition of q-gap, a and b belong to different q-gaps, indeed there exists $p, q \in \mathcal{M}$ such that $(a, p) \prec (b, q)$. Of course this contradicts the assumption that on $B \times \mathcal{M}$, $\prec$ is empty.

To prove the theorem, suppose that the set of pairs of elements in $B \times \mathcal{M}$ for which strict preference holds is nonempty. By lemma (4.10) there exists a subset B_h of B with $|B_h| = h \leqslant |A|$ and an enumeration of its elements $\{a_1, \ldots, a_h\}$ such that if $g_j := g_{a_j}$ and $l_j := l_{a_j}$, then

$$g_1 \prec g_2 \prec l_1 \precsim g_3 \prec l_2 \precsim g_4 \prec \cdots \prec g_{h-1} \prec l_{h-2} \precsim g_h \prec l_{h-1} \prec l_h.$$

I first show that the conditions that constitute a vNM space imply part (i) of the present theorem. By lemma (4.1) the exists a mixture preserving function $U_1 : \{a_1\} \times \mathcal{M} \to \mathcal{R}$ such that for all $p, q \in \mathcal{M}$

$$(a_1, p) \precsim (a_1, q) \quad \Leftrightarrow \quad U_1(a_1, p) \leqslant U_1(a_1, q)$$

By compactness of $\mathcal{M}$, for some $\beta_1, \gamma_1 \in \mathcal{R}$ I have $U_1(a_1, \mathcal{M}) = [\beta_1, \gamma_1]$ with $U_1(g_1) = U_1(a_1, \underline{p})$ and $U_1(l_1) = U_1(a_1, \overline{p})$ for some $\underline{p}, \overline{p} \in \mathcal{M}$.

Since $g_1 \prec g_2 \prec l_1$, lemma (4.3) implies that there exists unique $0 < \lambda < 1$, and $p(\lambda) := \underline{p}\lambda\overline{p}$ such that $(a_1, p(\lambda)) \sim g_2$. Similarly, let $g_2 = (a_2, \underline{q})$ and $l_2 = (a_2, \overline{q})$, then there exists unique $0 < \nu < 1$ such that $l_1 \sim (a_2, q(\nu))$, where $q(\nu) := \underline{q}\nu\overline{q}$.

By lemma (4.1), the vNM ordered set defined by projecting preferences onto the set $\{a_2\} \times \mathcal{M}$, has a mixture preserving representation $U_2' : \{a_2\} \times \mathcal{M} \to \mathcal{R}$. Thus, with a view to leaving the image of U_1 unchanged in my construction of a mixture preserving representation on the projection of preferences onto $\{a_1, a_2\} \times \mathcal{M}$, I recall lemma (4.2) states that: for any interval $[\beta_2, \gamma_2]$ in $\mathcal{R}$, there exists unique $\theta_2, \kappa_2 \in \mathcal{R}$, $\theta_2 > 0$ such that

$$[\theta_2 U_2'(g_2) + \kappa_2,\ \theta_2 U_2'(l_2) + \kappa_2] = [\beta_2, \gamma_2].$$

The content of two preceding paragraphs suggests that I should choose β_2 and γ_2 to satisfy $\beta_2 = \lambda\beta_1 + (1-\lambda)\gamma_1$ and $\gamma_1 = \nu\beta_2 + (1-\nu)\gamma_2$. Substituting for β_2 in the second of these two equations, I obtain $\gamma_2 = \frac{1}{1-\nu}(\gamma_1 - \nu\beta_2)$. Let

$U_2 : \{a_1, a_2\} \times \mathcal{M} \to \mathcal{R}$ be defined as

$$U_2(x) = \begin{cases} U_1(x) & \text{if } x \in \{a_1\} \times \mathcal{M} \\ \theta_2 U_2'(x) + \kappa_2 & \text{if } x \in \{a_2\} \times \mathcal{M} \end{cases}$$

U_2 is clearly mixture preserving and its image is the interval $[\beta_1, \gamma_2]$. In order to show that it is also a representation, it suffices to check pairs x and y where $x := (a_1, p')$ and $y := (a_2, q')$ for some $p', q' \in \mathcal{M}$.

First consider the more straightforward cases, that is where $x \prec g_2 \precsim y$ and $x \precsim l_1 \prec y$. For the former, since

$$U_2(g_2) = \beta_2 = \lambda\beta_1 + (1-\lambda)\gamma_1 = U_2(a_1, p(\lambda)),$$

the fact that $x = (a_1, p') \prec g_2 \sim (a_1, p(\lambda))$ together with theorem 4 and 6 of [HM] imply that there is a unique $\lambda_x < \lambda$ such that $x \sim (a_1, p(\lambda_x)$. In this case,

$$U_2(x) = U_2(a_1, p(\lambda_x)) = \lambda_x\beta_1 + (1-\lambda_x)\gamma_1 < \beta_2 \leqslant U_2(y),$$

and so I conclude that whenever $x \prec g_2 \precsim y$

$$x \precsim y \quad \Leftrightarrow \quad U_2(x) \leqslant U_2(y),$$

as required. (Note that in making this statement I have made use of the fact that in this case there are no pairs x, y such that $y \precsim x$.) For the set of pairs x, y where $x \precsim l_1 \prec y$ the proof that U_2 is a representation on such pairs is the same.

Now consider the remaining case where $x = (a_1, p')$ and $y = (a_2, q')$ for some $p', q' \in \mathcal{M}$ and $g_2 \precsim x, y \precsim l_1$. I recall that

$$(a_2, \underline{q}) = g_2 \sim (a_1, p(\lambda)) \quad \text{and} \quad (a_1, \overline{p}) = l_1 \sim (a_2, q(\nu)),$$

and once again by theorem 6 of [HM], there exists a unique $0 \leqslant \mu \leqslant 1$ with $p' \sim_{a_1} p(\lambda)\mu\overline{p}$ such that $x \sim \big(a_1, p(\lambda)\mu\overline{p}\big)$. By lemma (4.11) therefore

$$x \sim \big(a_1, p(\lambda)\mu\overline{p}\big) \sim \big(a_2, \underline{q}\mu q(\nu)\big).$$

Now suppose that $x \sim y$. In this case, transitivity implies that $y \sim \big(a_2, \underline{q}\mu q(\nu)\big)$. Then since $U_2(a_1, p(\lambda)) = \beta_2 = U_2(a_2, \underline{q})$ and $U_2(1, \overline{p}) = \gamma_1 = U_2(a_2, q(\nu))$ and U_2 is mixture preserving (whenever the mixture operation is defined) I have

$$U_2(a_1, p') = U_2(a_1, p(\lambda)\mu\overline{p}) = \mu\beta_2 + (1-\mu)\gamma_1$$

and

$$U_2(a_2, q') = U_2(a_1, \underline{q}\mu q(\nu)) = \mu\beta_2 + (1-\mu)\gamma_1$$

as required.

The remaining cases, where $x \prec y$ and $y \prec x$ follow by virtue of the following facts: by theorem 6 of [HM], I may find two unique values $0 < \mu_x, \mu_y < 1$ such that $p' \sim_{a_1} p(\lambda)\mu_x\overline{p}$ and $q' \sim_{a_2} \underline{q}\mu_y q(\nu)$; by lemma (4.11) I know that $(a, p(\lambda)\mu\overline{p}) \sim (a_2, \underline{q}\mu q(\nu))$ for all $0 \leqslant \mu \leqslant 1$; and by theorem 4 of [HM] I have

$\mu_x < \mu_y$ if and only if both

$$\underline{q}\mu_x q(\nu) \prec_{a_2} \underline{q}\mu_y q(\nu) \quad \text{and} \quad p(\lambda)\mu_x\overline{p} \prec_{a_1} p(\lambda)\mu_y\overline{p}.$$

Thus far we've seen that U_2 is a mixture preserving representation of the projection of preferences $\precsim$ onto the subset $\{a_1, a_2\} \times \mathcal{M}$ of the strict cover $B_h \times \mathcal{M}$. In order to extend U_2 to the rest of this strict cover, I proceed by induction and the proof follows by precisely the same argument as above. The resulting representation is a function $U_h : B_h \times \mathcal{M} \to [\beta_1, \gamma_h]$ that is a standard mixture preserving representation on each of the sets $\{a_j\} \times \mathcal{M}$ and an extended mixture preserving representation on the whole set.

In order to extend U_h to the rest of the quasi-component $B \times \mathcal{M}$, note that for all $b \in B \backslash B_h$, $g_1 \precsim g_b \precsim l_b \precsim l_h$. In fact, because a strict cover is minimal, I conclude that $g_j \precsim g_b \precsim l_b \precsim l_{j+1}$ for some $j \in \{1, \ldots, j-1\}$. (See the proof of lemma (4.10) for the construction of a strict cover.) Thus for some $x, y \in \{a_j, a_{j+1}\} \times \mathcal{M}$ I have $x \sim g_b$ and $y \sim l_b$.

Now by the proof that U_2 is a representation of the projection of preferences onto $\{a_1, a_2\} \times \mathcal{M}$, and the fact that my construction of U_h is recursive in j, I know that for each j, U_h restricted to $\{a_j, a_{j+1}\} \times \mathcal{M}$ will have image $[\beta_j, \gamma_{j+1}]$ where $\beta_j < \gamma_{j+1}$. If I choose $\beta_b = U_h(x)$ and $\gamma_b = U_h(y)$ then by lemma (4.2) or remark (4.1) given any representation U_b of the projection of preferences onto $\{b\} \times \mathcal{M}$ there exists a positive affine transformation $\theta_b > 0$, $\kappa_b \in \mathcal{R}$ that is geometrically unique and satisfies

$$[\beta_b, \gamma_b] = [\theta_b U_b(g_b) + \kappa_b, \theta_b U_b(l_b) + \kappa_b].$$

Indeed because this is true for all $b \in B \backslash B_h$, the proof that

$$U_B(x) := \begin{cases} U_h(x) & \text{if } x \in B_h \times \mathcal{M} \\ \theta_b U_b(x) + \kappa_b & \text{if } x \in \{b\} \times \mathcal{M},\ b \in B \backslash B_h \end{cases}$$

is a representation for the projection of preferences onto the quasi-component $B \times \mathcal{M}$ follows by the same techniques that I have used in showing that U_2 is a representation.

It remains to be shown that the fact that $(A \times \mathcal{M}, \precsim)$ is a vNM ordered space implies part (ii) of the present theorem. That is, if V is any other mixture preserving representation of there exists a single positive affine transformation $\theta > 0$, $\kappa \in \mathcal{R}$ such that $V = \theta U + \kappa$. If $V(B \times \mathcal{M}) = [\pi, \rho]$ for some $\pi < \rho$ in $\mathcal{R}$, then by the proof of lemma (4.2), I let θ satisfy $\theta(\gamma_h - \beta_1) = \rho$ and κ satisfy $\theta\beta_1 + \kappa = \pi$. I recall that in the construction of the image of U_h, the representation of the strict cover of $B \times \mathcal{M}$, only β_1 and γ_1 were free variables. The degrees of freedom associated with every other representation $\widetilde{U}_j$ of $\{a_j\} \times \mathcal{M}$ were used to obtain the unique transformation $\theta_j \widetilde{U} + \kappa_j$ with image $[\beta_j, \gamma_j]$, where for each j, $\beta_j = \lambda_j \beta_{j-1} + (1 - \lambda_j)\gamma_{j-1}$ and $\gamma_j = \frac{1}{1-\nu_j}(\gamma_{j-1} - \nu_j \beta_j)$ and where the λ_j and ν_j are uniquely determined by preferences. As such it suffices to check that U and V agree on $\{a_1\} \times \mathcal{M}$. This of course directly follows by lemma (4.2).

This complete the proof that the conditions that constitute an extended vNM ordered space imply (i) and (ii) of the present theorem. In the opposite direction, conditions (O), (C'ty) and (I) are standard and therefore omitted. The necessity of (CB) is seen by noting that if it fails, that is there exists $a, b \in B$, $p, q, p', q' \in \mathcal{M}$ with $(a, p) \sim (b, q)$ and $(a, p') \sim (b, q')$ but

$$(a, p\tfrac{1}{2}p') \nsim (b, q\tfrac{1}{2}q').$$

Without loss of generality, suppose that $(a, p) \prec (b, q)$, then for any mixture preserving representation U of such preferences $\beta = U(a, p) = U(b, q)$ and $\gamma = U(a, p') = U(b, q')$ for some $\beta < \gamma$ in $\mathcal{R}$. However, the fact that U is mixture preserving on each of $\{a\} \times \mathcal{M}$ and $\{b\} \times \mathcal{M}$ implies

$$\begin{aligned} U(p\mu p') &= \mu\beta + (1 - \mu)\gamma \\ &= U(b, q\mu q') \end{aligned}$$

for all $0 < \mu < 1$; which clearly implies the desired contradiction.

The necessity of the assumption that $B \times \mathcal{M}$ is a quasi-component of $(A \times \mathcal{M}, \precsim$) follows from the fact that if it is not then there exists at least two quasi-components $B' \times \mathcal{M}$ and $B'' \times \mathcal{M}$ whose union is $B \times \mathcal{M}$. I will obtain a contradiction for the case where there are only two q-components. Let g', g'', l' and l'' be the respective glbs and lubs of B' and B''. It will suffice to consider the cases where they are components, that is $g' \prec l' \prec g'' \prec l''$, and where they are quasi-components but not components, that is where $g' \prec l' \sim g'' \prec l''$.

First consider the case where $l' \prec g''$. Let U be a mixture preserving representation of $B \times \mathcal{M}$ with $U(g') = \beta'$, $U(l') = \gamma'$, $U(g'') = \beta''$ and $U(l'') = \gamma''$. Now let $[\pi', \rho']$ and $[\pi'', \rho'']$ be any pair of intervals such that $\pi' < \rho' < \pi'' < \rho''$. It is clear that in general I will need two different positive affine transformations $\theta', \theta'' > 0$ and $\kappa', \kappa'' \in \mathcal{R}$ are needed to shift and rescale the intervals $[\beta', \gamma']$ and $[\beta'', \gamma'']$ so that $[\theta'\beta' + \kappa', \theta'\gamma' + \kappa'] = [\pi', \rho']$ and $[\theta''\beta'' + \kappa'', \theta''\gamma'' + \kappa''] = [\pi'', \rho'']$. However it is also clear that any mixture preserving function that maps $B' \times \mathcal{M}$ into $[\pi', \rho']$ and $B'' \times \mathcal{M}$ into $[\pi'', \rho'']$ and which is representation on each of these quasi-components is also a representation on $B \times \mathcal{M}$. This implies that (ii) is violated. So that in this case, the fact that $B \times \mathcal{M}$ is a quasi-component is necessary.

The case where $g' \prec l' \sim g'' \prec l''$ follows by an identical argument and is therefore omitted. This completes the proof of the theorem. $\square$

Theorem 4.2 (Mixture preserving representation)**.**
Let $\mathcal{M}$ be a compact mixture space and A a finite set. Then $(A \times \mathcal{M}, \precsim)$ is an extended *vNM ordered space if and only if both the following conditions hold.*

(1) There exists a mixture preserving function $U : A \times \mathcal{M} \to \mathcal{R}$ such that for all $x, y \in A \times \mathcal{M}$

$$x \precsim y \quad \Leftrightarrow \quad U(x) \leqslant U(y).$$

(2) the number Q of quasi-components[9] *has the property that if V is any*

[9]For those who have not read the construction of quasi-components above, these should be interpreted as corresponding to maximal intervals in the image of the utility function in $\mathcal{R}$ that are non-overlapping except, possibly, at the endpoints. It is this slightly messy property that leads the construction to be involved even though the concept is straightforward.

other function satisfying (1), then for each $a \in A$

$$V(a,\cdot) = \theta_i U(a,\cdot) + \kappa_i$$

for some $i = 1,\ldots,Q$, $\kappa_i \in \mathcal{R}$ and $\theta_i > 0$.

Proof. ($\Rightarrow$ (1)) Since $A \times \mathcal{M}$ can be written as the union of an arbitrary enumeration of its q-components $A_i \times \mathcal{M}$, $i = 1,\ldots,Q \leqslant |A|$. (Recall that we are implicitly referring to distinct q-components.) Each q-component, by theorem (4.1), has a mixture preserving representation $\widetilde{U}^i : \{a_i\} \times \mathcal{M} \to [\beta_i', \gamma_i']$ for some $\beta_i' \leqslant \gamma_i'$ in $\mathcal{R}$. I proceed by rearranging the intervals $[\beta_i', \gamma_i']$, to match the order $\precsim$.[10]

The proof then follows by lemma (4.2) as we are free to take the necessary positive affine transformations of the representations $\widetilde{U}^i$ such that the transformed representation U^i has image equal to the desired interval $[\beta_i, \gamma_i]$. My mixture preserving representation will be the function that coincides with each of the functions U^i for each q-component $A_i \times \mathcal{M}$.

Let $[\beta_1, \gamma_1] = [0, 1]$ if $\beta_1' < \gamma_1'$ and $\{0\}$ otherwise. By induction, for the i_{th} interval in the enumeration, consider the following cases:

- $l_i \prec g_m$ for all $m \leqslant i-1$: if so let $[\beta_i, \gamma_i] = [-2^i - 1, -2^i]$ if $\beta_i' < \gamma_i'$ and $\{-2^i\}$ otherwise;
- $l_m \prec g_i$ for all $m \leqslant i-1$: if so let $[\beta_i, \gamma_i] = [2^i, 2^i + 1]$ if $\beta_i' < \gamma_i'$ and $\{2^i\}$ otherwise;

[10] This method of proof allows for a countable set A, although my construction of q-components precludes the theorem from applying to that case.

- $l_h \prec g_i \precsim l_i \prec g_k$ for some $1 \leqslant h, k \leqslant i-1$ and for every $m \leqslant i-1$ I have either $l_m \precsim l_h$ or $g_k \precsim g_m$.

If the last of these is true and $\beta_i' < \gamma_i'$, then let $\beta_i = \frac{1}{2}(\gamma_h + \beta_k - \frac{\beta_k - \gamma_h}{2^i})$ and $\gamma_i = \frac{1}{2}(\gamma_h + \beta_k + \frac{\beta_k - \gamma_h}{2^i})$ and let them both equal $\frac{1}{2}(\gamma_h + \beta_k)$ otherwise.

In the first of these cases, for each $i \geqslant 2$, then $\gamma_i = -2^i < -2^{i-1} - 1$ where $-2^{i-1} - 1$ is a lower bound for β_m for every $m \leqslant i-1$. The second case is similar. For the third case, it is also clear that for all $i \geqslant 2$ $\gamma_h < \beta_i \leqslant \gamma_i < \beta_k$.

Now consider the remaining cases where the q-component $[g_i, l_i]$ shares an indifference set with either one or two q-components with respect to which it is distinct. Recall, that by lemma (4.8) in this case, we cannot have $g_i \sim l_i$.

- $l_i \sim g_h$, where $g_h \precsim g_m$ for all $m \leqslant i-1$: if so let $[\beta_i, \gamma_i] = [\beta_h - 1, \beta_h]$;
- $l_k \sim g_i$, where $l_m \precsim l_k$ for all $m \leqslant i-1$: if so let $[\beta_i, \gamma_i] = [\gamma_k, \gamma_k + 1]$;
- $l_h \precsim g_i \prec l_i \precsim g_k$ for some $1 \leqslant h, k \leqslant i-1$ and for every $m \leqslant i-1$ we have either $l_m \precsim l_h$ or $g_k \precsim g_m$.

If the last of these is true, then there are three sub-cases to consider. If $l_h \sim g_i \prec l_i \sim g_k$, then I simply set $[\beta_i, \gamma_i] = [\gamma_h, \beta_k]$. If $l_h \prec g_i \prec l_i \sim g_k$, then let $\gamma_i = \beta_k$ and $\beta_i = \beta_k - \frac{\beta_k - \gamma_h}{2^i}$. Similarly, if $l_h \sim g_i \prec l_i \prec g_k$, then I let $\beta_i = \gamma_h$ and $\gamma_i = \gamma_h + \frac{\beta_k - \gamma_h}{2^i}$. □

For the special case, where Δ is the set of probability distributions over fixed finite set, I obtain the following representation.

Theorem 4.3 (vNM for $A \times \Delta$)**.**

Let A and S be finite sets and let Δ be the set of probability measures on S. Then $(A \times \Delta, \precsim)$ is an extended *vNM ordered space if and only if both the following conditions hold.*

(1) There exists a function $U : A \times \Delta \to \mathcal{R}$, $(a,p) \mapsto U(a,p) = \sum_{s \in S} p_s u(a,s)$ such that for all $x, y \in A \times \Delta$

$$x \precsim y \quad \Leftrightarrow \quad U(x) \leqslant U(y).$$

(2) The number Q of quasi-components[11] satisfies the property that if V is any other function satisfying (1), then for each $a \in A$

$$V(a, \cdot) = \theta_i U(a, \cdot) + \kappa_i$$

for some $i = 1, \ldots, Q$, $\kappa_i \in \mathcal{R}$ and $\theta_i > 0$.

4.4 Applications

We will first discuss the relationship between the present work and that of [KS] and in the following two subsections, present applications of the above results to study the issue of state-dependence and the paradox of Allais (1953).

The following remark describes the relationship between condition (CB) and the corresponding condition of [KS].

Remark 4.4. *The following is the constrained independence condition of [KS].*

[11] We recall that these should be interpreted as corresponding to maximal intervals in the image of the utility function in $\mathcal{R}$ that are non-overlapping except, possibly, at the endpoints.

I emphasize that this is not assumed anywhere in this chapter, and is only presented so as to show that it is somewhat stronger than (CB).

Definition 4.11 (Constrained independence)**.** *For any $b, c \in A$, $p, q, p', q' \in \mathcal{M}$, and $\lambda \in [0,1]$ if $(b,p) \sim (c,q)$ then $(b,p') \precsim (c,q')$ if and only if*

$$(b, p\lambda p') \precsim (c, q\lambda q').$$

The fact that constrained independence implies condition (CB) follows immediately if we take $\lambda = \frac{1}{2}$, consider the fact that $\sim \subset \precsim$ and use the "only if" part of the statement. Moreover, it is clear that in the absence of condition (O), (CB) does not imply the "if" part of constrained independence, for (b,p') and (c,q') may be incomparable. We claim without proof, that in the presence of (O) and (C'ty) the conditions are equivalent.

The remaining conditions of [KS] serve only to ensure that the representation is state-independent and is particular to the relationship between the the set A and the mixture space $\mathcal{M}$ that they consider. That is to say, their paper is a special case, with additional conditions on preferences and structure on the space on which preferences are defined. Thus to my knowledge, this chapter constitutes the most general in its class, and is the natural generalization of vNM and [HM] to the setting where there are is more than one mixture space, or where the mixture operation is not everywhere defined.

4.4.1 Separating comparability from state-dependence

Consider the following change in interpretation of the space $A \times \mathcal{M}$ of the above theorems. Take A to be the product space

$$\prod_{s \in S} A_s$$

for some finite set of states S and, for each $s \in S$ a finite set of state-outcomes A_s. Then take $\mathcal{M}$ to be the set of probability distributions $\Delta(S)$ on S. Now as an element of $A \times \Delta(S)$ is of the form $(a,p) := \big((a_1, \ldots, a_{|S|}), (p_1, \ldots, p_{|S|})\big)$ which, in the case where the elements of A are also vectors in $\mathcal{R}^{|S|}$, may be rewritten as an inner product

$$\langle a, p \rangle = \sum_{s \in S} p_s \, a_s.$$

Or, in the usual lottery form where $\delta(a_s)$ is the lottery assigning probability one to the state-outcome a_s

$$\sum_{s \in S} \delta(a_s) \, p_s.$$

Now for each a in the finite set A, $a \times \Delta(S)$ is a mixture space, and so by theorem (4.3) preferences satisfy (O), (C'ty) and (CB), if and only if there exists a function $U : A \times \Delta(S) \to \mathcal{R}$, $(a,p) \mapsto U(a,p) = \sum_{s \in S} p_s u_s(a_s)$ such that for all $(b,q),(c,r) \in A \times \Delta$

$$(b,q) \precsim (c,r) \quad \Leftrightarrow \quad \sum_{s \in S} u_s(b_s) \, q_s \leqslant \sum_{s \in S} u_s(c_s) \, r_s \, .$$

Note that the means to which we obtain this may be easier to see if we write $u_s(a_s) = u(a_s, \delta_s)$, but it is standard in the literature to make this small abuse of notation. Now this is clearly a state dependent utility function. That is, for any pair of states $s, t \in S$, it may well be that $b_s = b_t = b_\cdot$, $c_s = c_t = c_\cdot$, $u_s(b_\cdot) < u_s(c_\cdot)$ and $u_t(c_\cdot) < u_t(b_\cdot)$ simultaneously hold. By contrast, theorem (4.3) implies that the uniqueness properties of the representation depend entirely upon preferences. Suppose that the extended vNM ordered set $(A \times \Delta(S), \precsim)$ defines a single component. In this case, any other representation of preferences can be rewritten as a positive affine transformation of U. Thus, in the presence of condition (CB), full comparability across states and state-dependence of preferences is possible.

Indeed, this is the property that ensures the representation of [KS] and Karni (2009) has the uniqueness form that it does. They obtain this by assuming a condition called "coordinate essentiality". This states that, for each state s, there exists $b, c \in A$ such that $(b, \delta_s) \prec (c, \delta_s)$. Whilst this condition is sufficient for the representation to be unique in the sense I have just described, contrary to what [KS] and Karni (2009) claim, it is certainly not necessary. The following example highlights this fact.

Example 4.4. *Suppose there are two states. Take A_1 to be a singleton.*[12] *Then preferences do not satisfy the condition in question. Now take A_2 to be of cardinality 2 and note that the cardinality of A is therefore also two. Now suppose that $(c, \delta_1) \equiv (b, \delta_1) \prec (b, \delta_2) \prec (c, \delta_2)$. Then by lemma (4.3) we know that there exists a unique $0 < \lambda < 1$ such that $(b, \delta_2) \sim (c, \delta_1 \lambda \delta_2)$. Since $(b, \delta_1) \prec (b, \delta_2)$. We conclude that there are two distinct indifference sets each containing elements from both the mixture spaces $\{b\} \times [0, 1]$ and $\{c\} \times [0, 1]$, and so, in the language of the present chapter, preferences give*

[12] We could also suppose that $(b, \delta_1) \sim (c, \delta_1)$ for all $b, c \in A$.

rise to a single component. That is, any representation of preferences is unique up to a single positive affine transformation. Thus, there exist preferences that satisfy part (b) of theorem 2 in [KS], but not part (a) of the same theorem. The condition in question is, however, needed for the uniqueness of the representation, for without it, there is no guarantee that preferences are of the form $(c, \delta_2) \prec (c, \delta_1) \equiv (b, \delta_1) \prec (b, \delta_2)$*: in this case, preferences give rise to two quasi-components and, by the main theorem of this chapter, two positive affine transformations characterize the uniqueness of the representation, not one. Grant, Kajii, Polak and Safra (2010) present the correct form of the theorem in a different setting.*

The "certainty principle" of [KS] states that for all states s, when the decision-maker is certain of being in state s, what might happen in other states doesn't matter, that is, she reduces the comparison of (a, δ_s) and (b, δ_s) to the comparison of a_s and b_s. Whilst this is not necessary for the uniqueness discussion above, it is necessary and sufficient for there to be a state-independent utility representation of the form

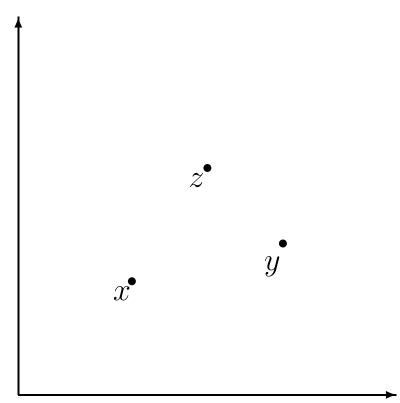
$$U : A \times \Delta(S) \to \mathcal{R}, \quad (a, p) \mapsto U(a, p) = \sum_{s \in S} p_s u(a_s),$$

where now, since $u(a_s, \delta_s) = u(a_s)$ for all s, we see that the function u depends only on the state through the state-outcome a_s.

The main message to take away from the present subsection is that, contrary to the common perception, state-dependence and uniqueness are independent concepts, the confusion seems to have arisen due to the almost universal use of a single mixture space.

4.4.2 Incompletely defined preferences and the Allais Paradox

By preferences that are incompletely defined we mean that there exist entire sets of lotteries, say, that the decision-maker has not even considered, but wherever preferences are defined they are complete. This is a special case of incomplete preferences which holds when there exists at least one pair of alternatives for which preferences have nothing to say, perhaps as a result of the decision-maker genuinely being unable to state preference in either direction. For instance, for incomplete preferences in general, it may be the case that x is worse than both y and z, whilst y and z are incomparable as in some ways y dominates z and in others z dominates y. The following diagram presents such a scenario for a standard (element by element) ordering of vectors in R^2.



Instead, by incompletely defined preferences we suppose that comparability is transitive. When two lotteries, x and y, are comparable, we mean that either x is preferred to y or y is preferred to x. By transitivity of the relation of of being comparable that if x is comparable to both y and z, then y is comparable with z. If we denote "comparable" by "$\leftrightarrow$" then this condition is summarized thus: for all x, y and z

$$\text{if} \quad x \leftrightarrow y \quad \text{and} \quad y \leftrightarrow z \quad \text{then} \quad x \leftrightarrow z.$$

This assumption ensures that the sets where preferences are complete are distinct from those where they are complete. By itself it is stronger than the usual transitivity of preference condition, but it is not strong enough to partition X into two sets, one containing comparable elements and the other elements for which preferences are not defined. There may be two or more sets within each of which all elements are comparable, but across which elements are not, or at least not comparable in the same way.

The concept at play is closely related to the approach of Schmeidler (1989). On p.576 he concedes that completeness is the most restrictive and imposing assumption of expected utility theory, but notes that "*One can view the weakening of the completeness assumption as a main contribution of all the other axioms.*" Indeed he then goes on to weaken the independence condition so that it applies only for a particular subclass of lotteries. The present application pertains to the same viewpoint.

Specifically, suppose that, for a fixed set of prizes such as $\{\$0, \$10, \$50\}$, a given decision-maker has been offered the choice between two pairs of lotteries over the prizes such as the pair $p = 0.01 \cdot \delta_0 + 0.99 \cdot \delta_{10}$ (win nothing with probability 0.01 and win \$10 with probability 0.99) and

$$q = 0.17 \cdot \delta_0 + 0.83 \cdot \delta_{50},$$

and the pair

$$r = 0.9 \cdot \delta_0 + 0.1 \cdot \delta_{50} \quad \text{and} \quad u = 0.11 \cdot \delta_0 + 0.89 \cdot \delta_{10},$$

where, once again, δ_x is the measure assigning probability one to outcome x. Now the chances are she will not have given much thought to the vast number of other possible lotteries over the set of prizes. However, she would almost certainly agree that both of p and q are strictly better than either r or u. In fact this may be so apparent that, when making such comparisons, the decision-maker may have no need for the high resolution scale of measurement the vNM model implies. After all, is it not the case that a person who is asked to state which of two rather different weights is heavier would have no need to ask if they could use scales before providing an answer?

By contrast, her decisions between p and q, on the one hand, and r and u on the other, are likely to require a good deal more consideration. Moreover, whilst for a given pair, such as p and q, mixtures nearby or in between may give rise to similar judgements, and be approximated by the vNM conditions, this need not hold globally. Unless more lotteries are placed before the decision-maker, thus allowing her to explore her own attitude to risk, her preferences may not even be defined globally. Near r and u, for instance, there may be another region where the model of vNM is a good approximation, but across the two regions, other than ordinal statements of the form "*anything in the region near p and q dominates anything in the region near r and u*, the decision-maker may be agnostic.

It seems reasonable that the utility function that characterizes such a decision-maker's preferences should reflect this reflect this asymmetry in the decision-making task across and within regions of the simplex. Furthermore, without further probing by the experimenter, say, the decision-maker may have no cause whatsoever to compare mixtures of elements in one region with those of the other, and if preferences are not defined there, then why should the utility function be?

In the example we have been considering we have so far left the preferences of the decision-maker unspecified. If however they took the form we see in figure (4.4.2) (see final page) we obtain an example of preferences that satisfy the Allais paradox, and which are also well described by the discussion above. If, as in the shaded regions of the diagram, we assume that each of these sets are convex, then they are mixture spaces and the model of this chapter provides a simple way of representing preferences with a family of mixture preserving utility functions (a simple generalization of an expected utility function that is defined below). Each member of this family being defined on one of the regions for which preferences are well defined. Moreover, these functions combine to define a single mixture preserving utility function on the union of the given regions.

On the complement of this union, preferences and hence the above utility function is simply not defined. However, if we extend each of the mixture preserving utility functions from their domain of definition to the whole simplex, each will take the form of a vNM expected utility function (also defined below). Together, over the entire simplex, these combine to form a multi-expected utility function that resembles that of Dubra, Maccheroni and Ok (2001).

The resulting representation may well fail to characterize preferences in the sense that, for some pair of lotteries p, q each of the vNM expected utility functions may happen to agree that p assigns greater utility than q, even though they both lie outside the regions where preferences are defined. Nonetheless, further research into finding appropriate conditions may provide a way of completing preferences for such pairs.

Another important difference between the representations is in the uniqueness properties the multi-expected utility functions possess. In the present model preferences may be such that each member of the family of utilities is numerically fully comparable with one or more of the others. That is, there may even be a single (multi-)utility scale. This may well be useful in applications.

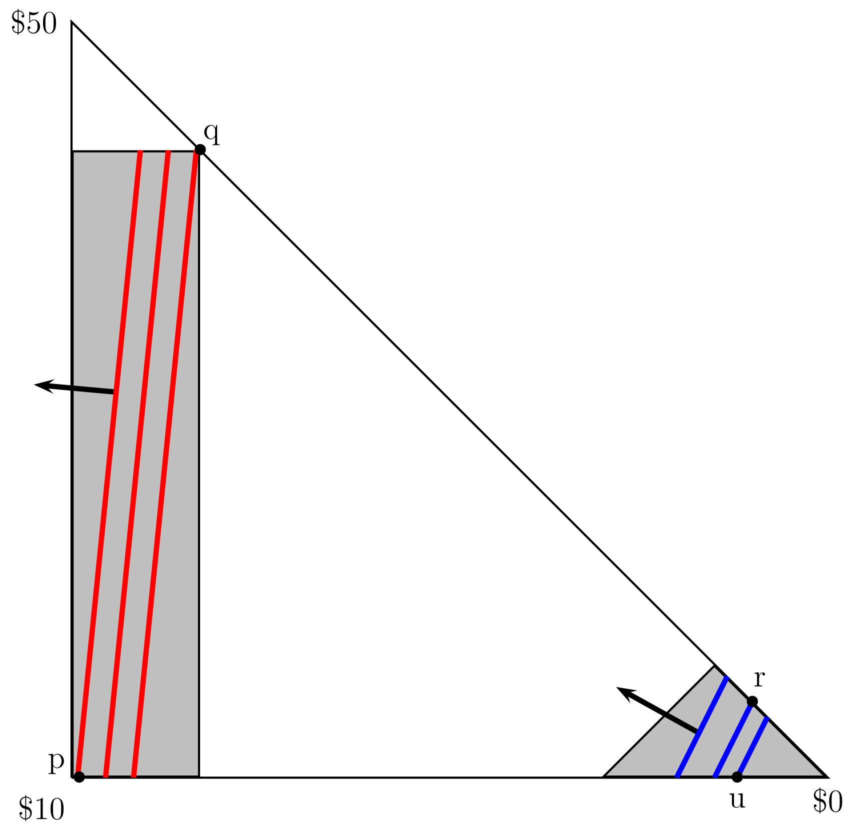


Figure 4.1: The Allais paradox and incomplete preferences. The decision-maker's preferences are such that the slope of the indifference curves in the triangular region containing δ_0 are shallower than those in the rectangular region containing p and q. If preferences satisfy the conditions for the main theorem of this paper over the union of these two regions, then the decision maker has a mixture preserving utility representation on this union. The representation is the restriction of two distinct expected utility functions to the respective regions. Since every element in the rectangle dominates every element in the triangle, preferences are lexicographic across regions and so comparing the utility value of an element in the rectangle with that of an element in the triangle is meaningless. Only ordinal statements are meaningful. By contrast, within each region, the representation is numerically meaningful in the sense of vNM and HM53.

Chapter 5

Summary and future research

In this thesis I have presented three approaches to dealing with preferences that vary with context. Whilst most of the results were presented in the canonical setup, where the space of contexts consists of probability distributions over some set of states, the task of transposing these results to different spaces is certainly feasible.

The first and most obvious transposition is to case-based decision theory, where the context space is the set of possible databases (memories), which is represented by $\mathbb{N}^n$: the product of n copies of the natural numbers. Clearly, there is an natural embedding in $\mathcal{R}^n_{\geqslant 0}$ of this space, and indeed, it is this property that Gilboa and Schmeidler (2001) employ in the proof of their representation theorem. When the context space is $\mathbb{N}^n$, the conditions on preferences that are analogous to the betweenness conditions of chapters 2 and 3, will be reminiscent of operations on cones as opposed to convex spaces. Thus, with the understanding that A is as usual the set of alternatives, condition (s-B) of chapter 3 would translate as follows.

Definition 5.1 (Additivity w.r.t. cases)**.** *Let δ_s be an element of $\mathbb{N}^n$ with s^{th} entry equal to 1 and all others equal to zero. For all j, δ_s in $\mathbb{N}^n$ and $a, b \in A$, if both $a \succ_j b$ and $a \succ_{\delta_s} b$, then $a \succ_{j+\delta_s} b$. The same condition holds for the relation $\succsim_{\cdot}$.*

The intuition for this kind of condition is similar to that of (s-B). Indeed in this setting, it is perhaps best to call it 'case monotonicity' due to the fact that $j \leqslant j + \delta_s$ for any j and any δ_s. This would be a weakening of the Gilboa and Schmeidler model since their 'combination' condition, addition is assumed to hold not only with respect to cases, but also any other database. That is, the decision-maker is assumed to be able to imagine what it would be like to concatenate any two databases, clearly this is a much more difficult exercise than simply imagining that another instance of the basic ingredients of memories were to occur: that is another case.

An additional condition that would substitute for the Archimedean one of Gilboa and Schmeidler (2001) would be the following.

Definition 5.2. *For every $j, \delta_s \in \mathbb{N}^n$ and every $a, b \in A$, if $a \succ_{\delta_s} b$, then there exists a natural number l, such that $a \succ_{j+l\delta_s} b$.*

If context (strict) preferences are also asymmetric and negatively transitive, then, in context space, preferences will resemble those in the following figure and preferences the objective is to show, using the techniques developed in chapter 3, that there is an ordinal representation with connected sets where no-strict preference holds between two alternatives and where star-convexity holds with respect to the elements δ_s as opposed to all databases j in $\mathbb{N}^n$.

For the model of chapter 2, where a cardinal representation characterizes preferences, the next step to weaken the diversity condition further. The way I propose to do this is to introduce conditional diversity. Roughly speaking it says that for any triple of alternatives, if there is a certain degree of diversity, then there is full diversity, but a lack of diversity is also allowed. This condition is a more substantial weakening Gilboa and Schmeidler's diversity condition, for it would allow, amongst other things, alternatives that are dominant for all contexts; thus allowing the model to encompass dominant strategy problems in games for instance.

For the model of chapter 4, where preferences were extended to contexts and the context space was a mixture space, the application to the Allais paradox needs to be expanded and formalized. In particular, to allow for collection of mixture spaces each of which is endogenously defined by preferences.

Finally, all of the above need to be extended to the multi-stage/decision-tree setting as well as the game-theoretic setting where the opponent may be strategic and it pays to be unpredictable.

Bibliography

[1] E. W. Adams. Elements of a theory of inexact measurement. *Philosophy of Science*, 32(3/4):205–228, 1965.

[2] M. Allais. Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine. *Econometrica*, 21(4):503–546, 1953.

[3] T. Amemiya. Qualitative response models: A survey. *Journal of Economic Literature*, 19(4):1483–1536, 1981.

[4] G. Ashkenazi and E. Lehrer. Relative utility. Preprint available at `www.math.tau.ac.il/ lehrer/Papers/relativeutil8.pdf`, 2001.

[5] Y. Azrieli. Axioms for Euclidean preferences with a valence dimension. *Journal of Mathematical Economics*, 47:545–553, 2011.

[6] S. Barberá, P. J. Hammond, and C. Seidl. *Handbook of Utility Theory: Extensions*. Handbook of Utility Theory. Kluwer Academic Publishers, 2004.

[7] T. F. Bewley. Knightian decision theory: Part I. *Cowles foundation working paper.* No. 807, Yale University, 1986.

[8] K. G. Binmore. *Rational decisions*. Gorman lectures in economics. Princeton University Press, 2009.

[9] C. Blackorby, D. Donaldson, and J.A Weymark. Social choice with interpersonal utility comparisons: a diagrammatic introduction. *International Economic Review*, 25(2):327–356, 1984.

[10] H. Bleichrodt. Reference-dependent utility with shifting reference points and incomplete preferences. *Journal of mathematical psychology*, 51:266–276, 2007.

[11] N. Bourbaki. *Elements of Mathematics: Algebra 1*, volume 2. Springer, 1998.

[12] D. Braziunas and C. Boutilier. Assessing regret-based preference elicitation with the utpref recommendation system. *Proceedings of the Eleventh ACM Conference on Electronic Commerce (EC'10)*, pages pp.219–228, 2010.

[13] C. P. Chambers and T. Hayashi. Choice and individual welfare. *Journal of economic theory*, 147(5):591–604, 2012.

[14] S. H. Chew. Axiomatic utility theories with the betweenness property. *Annals of Operations Research*, 19(1):273–298, 1989.

[15] C. d'Aspremont and L. Gevers. Social welfare functionals and interpersonal comparability. In K. Arrow, A. Sen, and K. Suzumura, editors, *Handbook of social choice and welfare*, volume 1, pages 459–541, 2002.

[16] M. H. A. Davis. *Markov models and optimization*. Monographs on statistics and applied probability. Chapman & Hall, 1993.

[17] G. Debreu. Representation of a preference ordering by a numerical function. In R. M. Thrall, C. H. Coombs, and R. L. Davis, editors, *Decision processes*, pages 159–165, 1954.

[18] G. Debreu. Topological methods in cardinal utility theory. *Cowles Foundation Discussion Papers*, 1959.

[19] G. Debreu. Continuity properties of paretian utility. *International Economic Review*, 5:285–293, 1964.

[20] E. Dekel. An axiomatic characterization of preferences under uncertainty: Weakening the independence axiom. *Journal of Economic Theory*, 40(2):304–318, 1986.

[21] J. H. Drèze. Les fondemonts logiques des l'utilite cardinale et de la probabilité subjective. *La Décision.* Paris: Colloques Internationaux du CNRS, 1961.

[22] J. Dubra, F. Maccheroni, and E. A. Ok. Expected utility theory without the completeness axiom. *Cowles Foundation Discussion Papers 1294*, 2001.

[23] S. Evans. Lecture notes from the 2010 probability at warwick workshop. Available at the University of Warwick Statistics Department P@W website., 2010.

[24] P. C. Fishburn. Semiorders and risky choices. *Journal of mathematical psychology*, 5:358–361, 1968.

[25] P. C. Fishburn. *Utility theory for decision making.* Publications in operations research. R. E. Krieger Pub. Co., 1979.

[26] P. C. Fishburn. Multilinear expected utility. *Mathematics of Operations Research*, 5(4):pp. 502–509, 1980.

[27] P. C. Fishburn. An axiomatic characterization of skew-symmetric bilinear functionals, with applications to utility theory. *Economics Letters*, 8(4):311–313, 1981.

[28] P. C. Fishburn and I. H. Lavalle. Context-dependent choice with nonlinear and nontransitive preferences. *Econometrica*, 56(5):1221–1239, 1988.

[29] I. Gilboa. *Theory of decision under uncertainty.* Econometric Society monographs. Cambridge University Press, 2009.

[30] I. Gilboa and D. Schmeidler. Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18:141–153, 1989.

[31] I. Gilboa and D. Schmeidler. Case-based decision theory. *The Quarterly Journal of Economics*, 110(3):605–639, 1995.

[32] I. Gilboa and D. Schmeidler. *A theory of case-based decisions.* Cambridge University Press, 2001.

[33] I. Gilboa and D. Schmeidler. A derivation of expected utility maximization in the context of a game. *Games and Economic Behavior*, 44(1):172 – 182, 2003.

[34] S. Grant, A. Kajii, B. Polak, and Z. Safra. Generalized utilitarianism and Harsanyi's impartial observer theorem. *Econometrica*, 78(6):1939–1971, 2010.

[35] V. Guillemin and A. Pollack. *Differential topology.* Mathematics Series. Prentice-Hall, 1974.

[36] P. R. Halmos. *Finite-dimensional vector spaces*. Springer-Verlag, New York, 1974.

[37] P. J. Hammond. Ex-ante and ex-post welfare optimality under uncertainty. *Economica*, 48(191):235–250, 1981.

[38] P. J. Hammond. Consequentialist foundations for expected utility. *Theory and Decision*, 25:25–78, 1988. 10.1007/BF00129168.

[39] P. J. Hammond. Consistent plans, consequentialism, and expected utility. *Econometrica*, 57(6):1445–1449, 1989.

[40] I. N. Herstein and J. Milnor. An axiomatic approach to measurable utility. *Econometrica*, 21(2):pp. 291–297, 1953.

[41] W. Hurewicz and H. Wallman. *Dimension theory*. Princeton mathematical series. Princeton University Press, revised 1948 edition, 1941.

[42] D. Kahneman and A. Tversky. Prospect theory: and analysis of decision under risk. *Econometrica*, 47(2):263–291, 1979.

[43] M. Kaneko. Epistemic logics and their game theoretic applications: Introduction. *Economic Theory*, 19(1):7–62, 2002.

[44] E. Karni. A theory of medical decision making under uncertainty. *Journal of Risk and Uncertainty*, 39(1):1–16, 2009.

[45] E. Karni and Z. Safra. An extension of a theorem of von Neumann and Morgenstern with an application to social choice theory. *Journal of Mathematical Economics*, 34(3):315–327, 2000.

[46] V. Köbberling and P. P. Wakker. Preference foundations for nonexpected utility: A generalized and simplified technique. *Mathematics of Operations Research*, 28(3):395–423, 2003.

[47] D. H. Krantz, R. D. Luce, Suppes P., and Tversky A. *Foundations of measurement, Vol.I: Additive and polynomial representations*. New York: Academic Press, 1971.

[48] D. M. Kreps. *Notes on the theory of choice*. Underground classics in economics. Westview Press, 1988.

[49] K. Kuratowski. *Topology*, volume 1 of *Topology*. Academic Press, 1968.

[50] K. Kuratowski. *Topology*, volume 2 of *Topology*. Academic Press, 1968.

[51] G. Loomes and R. Sugden. Regret theory: An alternative theory of rational choice under uncertainty. *The Economic Journal*, 92(368):805–824, 1982.

[52] A. Mas-Colell, M.D. Whinston, and J.R. Green. *Microeconomic theory*. Oxford University Press, 1995.

[53] D. L. McFadden. *Econometric analysis of qualitative response models. Handbook of Econometrics*, volume II, chapter 24. Elsevier Science Publishers BV, 1984.

[54] P. Mongin. A note on mixture sets in decision theory. *Decisions in Economics and Finance*, 24(1):59–69, 2001.

[55] S. Morris and T. Ui. Best response equivalence. *Games and Economic Behavior*, 49(2):260–287, 2004.

[56] J. R. Munkres. *Topology*. Prentice Hall, 2nd edition, 2000.

[57] R. B. Myerson. Axiomatic derivation of scoring rules without the ordering assumption. *Social choice and welfare*, 12:59–74, 1995.

[58] Y. Nakamura. Expected utility with an interval ordered structure. *Journal of mathematical psychology*, 32:298–312, 1988.

[59] P. Orlik. *Introduction to arrangements*. Regional conference series in mathematics. Published for the Conference Board of the Mathematical Sciences by the American Mathematical Society, 1989.

[60] K. W. S. Roberts. Interpersonal comparability and social choice theory. *The Review of Economic Studies*, 47(2):421–439, 1980.

[61] E. T. Rolls and G. Deco. *The Noisy Brain: Stochastic Dynamics as a Principle of Brain Function*. Oxford University Press, Oxford, 2010.

[62] A. Rubinstein. *Economics and Language*. Cambridge Univ. Press, 2000.

[63] P. Samuelson. A note on the pure theory of consumers' behaviour. *Economica*, 5:61–71, 1938.

[64] L. J. Savage. *The foundations of statistics*. Dover Publications, 1972.

[65] D. Schmeidler. Subjective probability and expected utility without additivity. *Econometrica*, 57(3):571–587, 1989.

[66] U. Schmidt. Reference dependence in cumulative prospect theory. *Journal of mathematical psychology*, 47:122–131, 2003.

[67] D. Scott and P. Suppes. Foundational aspects of theories of measurement. *The Journal of Symbolic Logic*, 23(2):113–128, 1958.

[68] E. Shafir, I. Simonson, and A. Tversky. Reason-based choice. *Cognition*, 49(1-2):11–36, 1993.

[69] L. A. Steen and J. A. Seebach. *Counterexamples in Topology*. Hot, Rinehart and Winston, Inc., USA, 1970.

[70] G. Strang. *Linear algebra and its applications*. Harcourt, Brace, Jovanovich, Publishers, 1988.

[71] A. Tversky and I. Simonson. Context-dependent preferences. *Management Science*, 39(10):1179–1189, 1993.

[72] K. Vind. Independent preferences. *Journal of Mathematical Economics*, 20(1):119–135, 1991.

[73] K. Vind and B. Grodal. *Independence, Additivity, Uncertainty*. Springer, http://springer.de, 2003.

[74] J. von Neumann and O. Morgenstern. *Theory of games and economic behavior*. Princeton University Press, Princeton and Oxford, Sixtieth anniversary edition, 1944.

[75] R. Webster. *Convexity*. Oxford science publications. Oxford University Press, 1994.

[76] P. H. Young. Social choice scoring function. *SIAM journal of applied mathematics*, 28:824–838, 1975.

[77] J. Zabczyk. *Chance and decision: stochastic control in discrete time*. Scuola Normale Superiore, 1996.